

126578

From: Borin, Michael  
Sent: Wednesday, July 07, 2004 11:51 AM  
To: STIC-Biotech/ChemLib  
Subject: Search request: 10/081108

CRPG

Examiner: M.Borin  
AU: 1631  
Mailbox:2C70  
Office: Remsen 2A55  
Tel.: 20713

RE: 10/081108

Please search :

1. nucleic acids of SEQ ID NO: 1.
2. polypeptide SEQ ID NO: 2.

against the commercial protein and nucleic acid databases + interference search.

Thank you

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: *July 7, 2004*  
Date Completed: *July 7, 2004*  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

TYPE OF SEARCH:  
NA Sequences: \_\_\_\_\_  
AA Sequences: \_\_\_\_\_ /  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: *olyp kosp*  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

**This Page Blank (uspto)**



# STIC Search Report

## Biotech-Chem Library

STIC Database Tracking Number: 126578

**TO:** Michael Borin  
**Location:** REM-2A55&2C70  
**Art Unit:** 1631  
**Friday, July 09, 2004**

**Case Serial Number:** 10/081108

**From:** Edward Hart  
**Location:** Biotech-Chem Library  
**REM-1A55**  
**Phone:** 571-272-2512

[edward.hart@uspto.gov](mailto:edward.hart@uspto.gov)

### Search Notes

Examiner Borin,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Edward Hart

**This Page Blank (uspto)**

Page 6

RA	Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T., Sakaki Y., Hattori M., Omura S.	"Complete genome sequence and comparative analysis of the industrial microorganism Streptomyces avermitilis." ; Nat. Biotechnol. 21:526-531(2003).
RT	GO; GO:003700; P:transcription factor activity; IEA.	DR; GO:006355; P:regulation of transcription, DNA-dependent; IEA.
GN	SEQUENCE FROM N.A.	DR; GO:000547; HTH LysR.
OS	STRAIN=C57BL/6J; TISSUE=Embryonic stem cells;	DR; InterPro; IPR005119; LysR_Subst.
OC	Mus musculus (Mouse).	DR; PF00122; HTH 1.
OC	Mammalia; Eutheria; Rodentia; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognath; Muridae; Mus.	DR; PF03466; LysR substrate; 1.
OX	NCBI_TAXID=10980;	DR; PROSITE; PS00044; HTH_LYSR_FAMILY; 1.
RN	SEQUENCE FROM N.A.	KW; Complete proteome.
RP	STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;	SQ; SEQUENCE 317 AA; 33592 MW; 9F6FBF58EC301AD1 CRC64;
RC	LINE=108560; Pubmed=1127851;	Query Match 26.1%; Score 55.5%; DB 16; Length 317;
RX	Kawai J., Shigenawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Arakawa T., Harr A., Fukunishi Y., Konno H., Adachi J., Fukuda S., Aizawa K., Iwasa M., Nishi K., Miyosawa H., Kondo S., Yamanaoka I., Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R., Kadota K., Matsuda H., Aishburner M., Battalov S., Casavant T., Fleischmann W., Gaasterland T., Gissi C., King B., Kohchiwa H., Kuehl P., Lewis S., Mataeu Y., Nakaido I., Pesole G., Quackenbush J., Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okuda T., Furuno M., Aono H., Baldarelli R., Barsh G., Blaize J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fletcher C., Fujita M., Garibaldi M., Gustincic S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P., Nordheim P., Ring B., Rungwald M., Rodriguez I., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Seiya T., Shiba Y., Storch K.-F., Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L., Wynnshaw-Boriss A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S., Hayashizaki Y.; "Functional annotation of a full-length mouse cDNA collection." ; Nature 409:685-690(2001).	Best Local Similarity 46.3%; Pred. No. 21; Indels 5; Gaps 1; Matches 19; Conservative 6; Mismatches 11; Indels 5; Gaps 1.
DR	Nature 409:685-690(2001).	RESULT 15
DR	AK010291; BAB26827.1; -.	Q9JNDS PRELIMINARY; PRT; 359 AA.
MGD	MGI:1889806; Tslpr.	AC; Q9JMD5; PRELIMINARY;
NON TER	1	DT 01-OCT-2000 (TREMBLrel. 15, Created)
FT	NON TER 126 AA; 13321 MW; BD387D755A8D9DF6 CRC64;	DT 01-OCT-2000 (TREMBLrel. 25, Last annotation update)
SQ	SEQUENCE 21 KEESPVY--SWRLPEDGTALCFI 42 :  :     :  :  :; 68 EEDDLIAKTRVPEPDGTSLCTV 92	Db 108 AadLVGAAALPGIL-HLAQZEPGVFSRLRAELEAGTAL 146
RP	SEQUENCE FROM N.A.	Cytochrome receptor delta1.
RC	LINE=1127851;	TSLPR.
RX	Q9JNDS PRELIMINARY; PRT; 359 AA.	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognath; Muridae; Mus.	OC
OC	NCBI_TAXID=10090;	NCBI_TAXID=10090;
RX	SEQUENCE FROM N.A.	RN SEQUENCE FROM N.A.
DR	LINE=20197866; PubMed=10733486;	DR LINE=20197866; PubMed=10733486;
DR	Q9JNDS PRELIMINARY; PRT; 359 AA.	DR PIR; JCT780; JCT780;
MGD	Fuji K.-, Nosaka T., Kojima T., Kawashima T., Yahata T., Nishimura T., Kitamura T.; "Molecular cloning of a novel type I cytokine receptor similar to the common gamma chain." ; Blood 95:2204-2211(2000).	DR MGD; MGI:1889506; Tslpr.
NON TER	RA Copeland N.G., Gilbert D.J., Jenkins N.A., Yamamoto K., Nishimura T., RA Kitamura T.; "Molecular cloning of a novel type I cytokine receptor similar to the common gamma chain." ; Blood 95:2204-2211(2000).	DR GO:0004872; P:receptor activity; IEA.
FT	RA SMART; SM00060; FN3; 1.	DR InterPro; IPR003961; FN3; 1.
SQ	SEQUENCE 21 KEESPVY--SWRLPEDGTALCFI 42 :  :     :  :  :; 68 EEDDLIAKTRVPEPDGTSLCTV 92	DR SMART; SM00060; FN3; 1.
RP	SEQUENCE FROM N.A.	KW; Receptor.
RC	LINE=21477403; PubMed=11572948;	SQ; SEQUENCE 359 AA; 37792 MW; 37FC8FF507B0E42B CRC44;
RX	Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C., Shinose M., Takahashi Y., Kurihara H., Nakazawa H., Osonee T., Kikuchi H., Shiba T., Sakaki Y., Hattori M.; "Genome sequence of an industrial microorganism Streptomyces avermitilis: deducing the ability of producing secondary metabolites." ; Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).	Query Match 26.1%; Score 55.5%; DB 11; Length 359;
OC	NCBI_TAXID=33903;	Best Local Similarity 44.0%; Pred. No. 24; Indels 4; Mismatches 7;
OX	SEQUENCE FROM N.A.	Db 301 EEDDLIAKTRVPEPDGTSLCTV 325
RN	STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;	Search completed: July 7, 2004, 17:09:07
RP	LINE=21477403; PubMed=11572948;	Job time : 48 secs
RP	Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C., Shinose M., Takahashi Y., Kurihara H., Nakazawa H., Osonee T., Kikuchi H., Shiba T., Sakaki Y., Hattori M.; "Genome sequence of an industrial microorganism Streptomyces avermitilis: deducing the ability of producing secondary metabolites." ; Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).	[2]
RX	SEQUENCE FROM N.A.	SEQUENCE FROM N.A.
RP	STRAIN=MA-680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;	SEQUENCE FROM N.A.
RP	LINE=2260316; PubMed=1262562;	SEQUENCE FROM N.A.
RX	LINE=2260316; PubMed=1262562;	SEQUENCE FROM N.A.







DE Hypothetical protein.  
 OS Homo sapiens (Human). Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1] \_  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Skin;  
 RA Strausberg R.; Submitted (May-2001) to the EMBL/GenBank/DDBJ databases.  
 RL EMBL; BC00040; AAH08404; 1.  
 DR GO; GO:0005829; Cytosol; ISS.  
 DR GO; GO:0005624; C:membrane fraction; ISS.  
 DR GO; GO:001050; F:D-erythro-sphingosine kinase activity; ISS.  
 DR GO; GO:000287; F:magnesium ion binding; ISS.  
 DR GO; GO:0002742; P:intracellular signaling cascade; ISS.  
 DR GO; GO:0046521; P:sphingoid catabolism; ISS.  
 DR InterPro; IPR001395; Aldo/ket\_red.  
 DR InterPro; IPR001206; DAGKC.  
 PFam; PF00781; DAGKC; 1.  
 Prodrom; PD005043; DAGKC; 1.  
 SMART; SM00046; DAGKC; 1.  
 PROSITE; PS00063; ALDOKETO\_REDUCTASE\_3; 1.  
 KW Hypothetical Protein.  
 DR GO; GO:00046; DAGKC; 1.  
 DR GO; GO:0046521; P:sphingoid catabolism; ISS.  
 KW SEQUENCE 384 AA; 42474 MW; F82999FF306113B0 CRC64;  
 SQ

Query Match. Score 29.1%; Best Local Similarity 40.0%; Matches 16; Conservative 8; Mismatches 10; Indels 6; Gaps 1; Pred. No. 3.1; DB 4; Length 384;

Qy 3 ARAVFLAQLQQLQARLMKES-----PVVSMRLEPDG 36  
 Db 295 SRAMLLRLFLAMERGRHMEYECPLVTVPPVAFRLEPDG 334

RESULT 3

Q96GKL ID Q96GKL PRELIMINARY; PRT; 398 AA.  
 AC Q96GKL; TRMBlrel. 19, Created  
 DT 01-DEC-2001 (TRMBlrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TRMBlrel. 25, Last annotation update)

DE Hypothetical protein.  
 OS Homo sapiens (Human).  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1] \_  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RA Strausberg R.; Submitted (Jun-2001) to the EMBL/GenBank/DDBJ databases.  
 RL EMBL; BC009419; AAH09419; 1.  
 DR GO; GO:0005829; Cytosol; ISS.  
 DR GO; GO:0005624; C:membrane fraction; ISS.  
 DR GO; GO:0017050; F:D-erythro-sphingosine kinase activity; ISS.  
 DR GO; GO:000287; F:magnesium ion binding; ISS.  
 DR GO; GO:0007242; P:intracellular signaling cascade; ISS.  
 DR GO; GO:0046521; P:sphingoid catabolism; ISS.  
 DR InterPro; IPR001395; Aldo/ket\_red.  
 DR InterPro; IPR001206; DAGKC.  
 PFam; PF00781; DAGKC; 1.  
 Prodrom; PD005043; DAGKC; 1.  
 SMART; SM00046; DAGKC; 1.  
 PROSITE; PS00063; ALDOKETO\_REDUCTASE\_3; 1.

KW Kinase.  
 SQ SEQUENCE 470 AA; 51084 MW; 5172B9A38C7CC17 CRC64;  
 Query Match. Score 29.1%; Best Local Similarity 40.0%; Matches 16; Conservative 8; Mismatches 10; Indels 6; Gaps 1; Pred. No. 3.8; DB 4; Length 470;

Qy 3 ARAVFLAQLQQLQARLMKES-----PVVSMRLEPDG 36  
 Db 381 SRAMLLRLFLAMERGRHMEYECPLVTVPPVAFRLEPDG 420

RESULT 5

Q87PBL ID Q87PBL PRELIMINARY; PRT; 668 AA.  
 AC Q87PBL; TRMBlrel. 24, Created  
 DT 01-JUN-2003 (TRMBlrel. 24, Last sequence update)

DE Tail-specific protease.  
 OS Vibrio parahaemolyticus.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibionales; Vibrionaceae; Vibrio.  
 OX NCBI\_TaxID=670;  
 RN [1] \_  
 RP SEQUENCE FROM N.A.  
 STRAIN=RIMD 21063 / Serotype O3:K6;  
 MEDLINE=22508154; PubMed=1262073;  
 RX Makino K., Osipima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K., Iijima Y., Naito M., Nakano M., Yamashita A., Kubota Y., Kimura S., Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;  
 RA "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism distinct from that of V. cholerae.";  
 RL Lancet 361:743-749 (2003); DR AP005078; BAD59861; -.  
 GO; GO:0008236; F:serine-type peptidase activity; IEA.  
 DR GO; GO:0007242; P:intracellular signaling cascade; IEA.  
 DR GO; GO:0007242; P:intracellular signaling cascade; IEA.

DE Hypothetical protein.  
 OS Homo sapiens (Human). Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1] \_  
 RP 309 SRAMLLRLFLAMERGRHMEYECPLVTVPPVAFRLEPDG 348  
 SQ Q8N632 PRELIMINARY; PRT; 470 AA.  
 AC Q8N632; TRMBlrel. 22, Created  
 DT 01-OCT-2002 (TRMBlrel. 22, Last sequence update)  
 DT 01-OCT-2003 (TRMBlrel. 25, Last annotation update)

DE Similar to sphingosine kinase.  
 OS Homo sapiens (Human).  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1] \_  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Blood, and Skin;  
 RA Strausberg R.; Submitted (May-2002) to the EMBL/GenBank/DDBJ databases.  
 RL EMBL; BC0010553; AAH30553; 1; -  
 DR EMBL; BC014439; AAH14439; 1.  
 DR GO; GO:0005829; C:cytosol; ISS.  
 DR GO; GO:0005624; C:membrane fraction; ISS.  
 DR GO; GO:0017050; F:D-erythro-sphingosine kinase activity; ISS.  
 DR GO; GO:000287; F:magnesium ion binding; ISS; cascade; ISS.  
 DR InterPro; IPR001395; Aldo/ket\_red.  
 DR InterPro; IPR001206; DAGKC.  
 PFam; PF00781; DAGKC; 1.  
 Prodrom; PD005043; DAGKC; 1.  
 SMART; SM00046; DAGKC; 1.  
 PROSITE; PS00063; ALDOKETO\_REDUCTASE\_3; 1.

KW Kinase.  
 SQ SEQUENCE 470 AA; 51084 MW; 5172B9A38C7CC17 CRC64;

Query Match. Score 29.1%; Best Local Similarity 40.0%; Matches 16; Conservative 8; Mismatches 10; Indels 6; Gaps 1; Pred. No. 3.8; DB 4; Length 470;

Qy 3 ARAVFLAQLQQLQARLMKES-----PVVSMRLEPDG 36  
 Db 381 SRAMLLRLFLAMERGRHMEYECPLVTVPPVAFRLEPDG 420

RESULT 4

Q8N632 ID Q8N632 PRELIMINARY; PRT; 470 AA.  
 AC Q8N632; TRMBlrel. 22, Created  
 DT 01-OCT-2002 (TRMBlrel. 22, Last sequence update)

DE Similar to sphingosine kinase.  
 OS Homo sapiens (Human).  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1] \_  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Blood, and Skin;  
 RA Strausberg R.; Submitted (May-2002) to the EMBL/GenBank/DDBJ databases.  
 RL EMBL; BC0010553; AAH30553; 1; -  
 DR EMBL; BC014439; AAH14439; 1.  
 DR GO; GO:0005829; C:cytosol; ISS.  
 DR GO; GO:0005624; C:membrane fraction; ISS.  
 DR GO; GO:0017050; F:D-erythro-sphingosine kinase activity; ISS.  
 DR GO; GO:000287; F:magnesium ion binding; ISS; cascade; ISS.  
 DR InterPro; IPR001395; Aldo/ket\_red.  
 DR InterPro; IPR001206; DAGKC.  
 PFam; PF00781; DAGKC; 1.  
 Prodrom; PD005043; DAGKC; 1.  
 SMART; SM00046; DAGKC; 1.  
 PROSITE; PS00063; ALDOKETO\_REDUCTASE\_3; 1.

KW Kinase.  
 SQ SEQUENCE 470 AA; 51084 MW; 5172B9A38C7CC17 CRC64;

Query Match. Score 29.1%; Best Local Similarity 40.0%; Matches 16; Conservative 8; Mismatches 10; Indels 6; Gaps 1; Pred. No. 3.2; DB 4; Length 398;

Qy 3 ARAVFLAQLQQLQARLMKES-----PVVSMRLEPDG 36  
 Db 381 SRAMLLRLFLAMERGRHMEYECPLVTVPPVAFRLEPDG 420

ALIGMENTS						
<b>SPTREMBL 25:*</b>						
1: sp_archaea:*						
2: sp_bacteria:*						
3: sp_fungi:*						
4: sp_human:*						
5: sp_invertebrate:*						
6: sp_mammal:*						
7: sp_micr:*						
8: sp_organelle:*						
9: sp_phage:*						
10: sp_plant:*						
11: sp_ratod:*						
12: sp_virus:*						
13: sp_vertebrate:*						
14: sp_unchassified:*						
15: sp_rvirus:*						
16: sp_bacteriap:*						
17: sp_archeap:*						
<b>RESULT 1</b>						
Q9BTG7	ID	Q9BTG7	PRELIMINARY;	PRT;	290 AA.	
	AC	Q9BTG7;				
	DT	01-JUN-2001	(TREMBLrel.	17,	Created)	
	DT	01-JUN-2001	(TREMBLrel.	17,	Last sequence update)	
	DT	01-JUN-2003	(TREMBLrel.	24,	Last annotation update)	
	DE		Similar to spingocine kinase 1 (Fragment).			
	OS	Homo sapiens (Human).				
	OC	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	OC	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.				
	OX	NCBI_TaxID=9606;				
	[1]					
RN	RP	SEQUENCE FROM N.A.				
	RC	TISSUE=Skin;				
	RA	Strausberg R.,				
	RL	Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.				
	DR	EMBL; BC004112; AAH04112.1; -				
	DR	GO; GO:0016301; F:kinase activity; IEA.				
KW						
FT	NON_TER	1	1			
SQ	SEQUENCE	290 AA;	31998 MW;	A3B7A219DB52CS20 CRC64;		
<b>SEQUENCE FROM N.A.</b>						
<b>RESULTS</b>						
<b>Pred.</b> No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.						
<b>SUMMARIES</b>						
result	No.	Score	Match	Length	DB	ID
Description						
Q9btg7	homo sapien					
Q9hv8	homo sapien					
Q96gk1	homo sapien					
Q9ns32	homo sapien					
Q87pb1	vibrio para					
Q9zq6	salmonella					
Q9vyc1	drosophila					
Q89ux1	bradyrhizob					
Q7u9q3	synchococc					
Q7urb8	rhodopirell					
Q7wuds	drosophilida					
Q8vr91	drosophila					
Q9crj5	mus musculu					
Q82pz2	streptomyce					
Q9jd55	mus musculu					
Q93jhs	mus musculu					
Q96hv8	PRELIMINARY;					
Q96hv8	ID					
Q96hv8	DT	01-DEC-2001	(TREMBLrel.	19,	Created)	
Q96hv8	DT	01-OCT-2003	(TREMBLrel.	19,	Last sequence update)	
Q96hv8	DT	01-OCT-2003	(TREMBLrel.	25,	Last annotation update)	
<b>RESULT 2</b>						
Q96hv8	PRT;	384 AA.				

Copyright GenCore version 5.1.6  
(c) 1993 - 2004 Compugen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: July 8, 2004, 23:35:12 ; Search time 6472 Seconds  
(without alignment)  
6723.796 Million cell updates/sec

Title: US-10-081-108-1

Perfect score: 1004

Sequence: 1 CGCCATTAGGGCTCTCGG.....ATCTTTACACTAAAAAGCC 1004

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:  
1: gb\_ba:\*

2: gb\_ntg:\*

3: gb\_in:\*

4: gb\_om:\*

5: gb\_nv:\*

6: gb\_pat:\*

7: gb\_ph:\*

8: gb\_Dl:\*

9: gb\_pr:\*

10: gb\_ro:\*

11: gd\_sts:\*

12: gb\_sy:\*

13: gb\_un:\*

14: gb\_vl:\*

15: em\_ba:\*

16: em\_fun:\*

17: em\_hum:\*

18: em\_ln:\*

19: em\_mu:\*

20: em\_om:\*

21: em\_xr:\*

22: em\_ov:\*

23: em\_dat:\*

24: em\_Db:\*

25: em\_Pl:\*

26: em\_ro:\*

27: em\_sts:\*

28: em\_ln:\*

29: em\_vl:\*

30: em\_htg\_hum:\*

31: em\_htg\_inv:\*

32: em\_htg\_other:\*

33: em\_htg\_mus:\*

34: em\_htg\_Pln:\*

35: em\_htg\_Rod:\*

36: em\_htg\_mam:\*

37: em\_htg\_rvt:\*

38: em\_sy:\*

39: em\_htg\_hum:\*

40: em\_htgo\_mus:\*

41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	1004	100.0	100.0	6	AR107991		AR107991 Sequence
2	1004	100.0	100.0	6	AR167384		AR167384 Sequence
3	1004	100.0	100.0	6	AR236412		AR236412 Sequence
4	1004	100.0	100.0	6	AR411457		AR411457 Sequence
5	1004	100.0	100.0	6	BD17601		BD17601 Method of
6	1004	100.0	100.0	6	HSU1980		HSU1980 Human B mel
7	966	103.2	103.2	6	AR104146		AR104146 Sequence
8	966	96.2	103.2	6	I28410		I28410 Sequence 1
9	966	103.2	103.2	6	I72215		I72215 Sequence 1
10	829.2	215.6	215.6	9	AF527550		AF527550 Homo sapi
11	563.4	106.8	106.8	9	AF527553		AF527553 Homo sapi
12	532.4	177.8	177.8	6	AC064811		AC064811 Homo sapi
13	529.2	139.0	139.0	6	AC133563		AC133563 Homo sapi
14	526	178.0	178.0	2	AC145113		AC145113 Homo sapi
15	524.4	1077.0	1077.0	2	AF155775		AF155775 Homo sapi
16	524.4	158.7	158.7	2	AC074294		AC074294 Homo sapi
17	524.4	182.0	182.0	8	AC140904		AC140904 Homo sapi
18	522.8	69.6	69.6	8	AC006175		AC006175 Homo sapi
19	522.8	189.3	189.3	6	AL161418		AL161418 Homo sapi
20	522.8	23.1	23.1	5	AL158811		AL158811 Homo sapi
21	521.2	19.0	19.0	2	AF254983		AF254983 Homo sapi
22	521.2	28.1	28.1	16	HS21C001		HS21C001 Homo sapi
23	467.8	148.6	148.6	6	AC10492		AC10492 Homo sapi
24	464.2	46.2	46.2	2	AF527552		AF527552 Homo sapi
25	448	44.6	44.6	9	AF527551		AF527551 Homo sapi
26	435.2	43.3	43.3	9	AF339516		AF339516 Homo sapi
27	433.6	18.4	18.4	9	AF339515		AF339515 Homo sapi
28	385	38.3	38.3	2	AF527554		AF527554 Homo sapi
29	266	26.5	26.5	9	AF339514		AF339514 Homo sapi
30	264.4	26.3	26.3	9	AF218570		AF218570 Homo sapi
31	215.4	21.5	21.5	9	AF499647		AF499647 Homo sapi
32	205.8	20.5	20.5	9	AL079472		AL079472 Homo sapi
33	205.8	18.0	18.0	2	AC140904		AC140904 Homo sapi
34	199.4	19.9	19.9	9	HSB117C8		HSB117C8 Human DNA
35	181.8	18.1	18.1	9	AC138036		AC138036 Homo sapi
36	181.8	12.9	12.9	9	E54646		E54646 Base sequen
37	181.8	14.2	14.2	9	AL163539		AL163539 Homo sapi
38	181.8	18.1	18.1	9	BX664725		BX664725 Homo sapi
39	181.8	23.9	23.9	9	BX088717		BX088717 Human DNA
40	179.4	17.9	17.9	9	AC008443		AC008443 Homo sapi
41	179.4	12.0	12.0	9	AL161615		AL161615 Homo sapi
42	179.4	17.9	17.9	9	HS16C2		HS16C2 Human
43	179.2	17.8	17.8	9	AL627334		AL627334 Homo sapi
44	179.2	19.0	19.0	9	AL050302		AL050302 Homo sapi
45	177.6	17.7	17.7	9	HSB62L20		HSB62L20

## ALIGNMENTS

RESULT 1	AR107991	LOCUS	Sequence 1 from patent US 6110694.	DEFINITION	1004 bp	DNA	lineair	PAT 14-FEB-2001
VERSION	AR107991	VERSION	AR107991	KEYWORDS				
SOURCE	Unknown.	ORGANISM	Unclassified.					
REFERENCE	1 (bases 1 to 1004)	AUTHORS	Boel,P., Wildmann,C., Boon-Falleur,T., van der Bruggen,P., Coulie,P. and Renaud,J.-C.					
TITLE	HLA-Cw*160/molecules							

JOURNAL Patent: US 6110694-A 1 29 AUG-2000;  
 FEATURES Location/Qualifiers  
 source 1 - 1004  
 /organism="unknown"  
 /mol\_type="unassigned DNA"

ORIGIN

Query Match Score 1004; DB 6; Length 1004;  
 Best Local Similarity 100.0%; Pred. No. 3.7e-223;  
 Matches 1004; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCCCAATTAGGGTTCGGTATCCTCGTAACTGTTAGAGAAC 60  
 Db 1 CGCCCAATTAGGGTTCGGTATCCTCGTAACTGTTAGAGAAC 60  
 Qy 61 AGGAAGGCCGAGCTGGAGCTGAGCCTGTAAACCGTGGCTCAGCTGATG 120  
 Db 61 AGGAAGGCCGAGCTGGAGCTGAGCCTGTAAACCGTGGCTCAGCTGATG 120  
 Qy 121 GTGGTGGCACACAGATGCCAGCTGGAGCTGGAGCTGAGCTGAGG 180  
 Db 121 GTGGTGGCACACAGATGCCAGCTGGAGCTGGAGCTGAGCTGAGG 180  
 Qy 181 AGTGGGGCGAGCTGGAGCTGAGATGCCGGCAAGGGCTTTCGGCATTTGTGCCAGC 240  
 Db 181 AGTGGGGCGAGCTGGAGCTGAGATGCCGGCAAGGGCTTTCGGATTTGTGCCAGC 240  
 Qy 241 TGCCTCAAGCCAGGTGATGAAGAGGACTCCCTGTTGACTGGTGGAGCTG 300  
 Db 241 TGCCTCAAGCCAGGTGATGAAGAGGACTCCCTGTTGACTGGTGGAGCTG 300  
 Qy 301 AAGACGGCACAGCTGTTGACTCTTGAGCTGGAGCTGGAGAGC 360  
 Db 301 AAGACGGCACAGCTGTTGACTCTTGAGCTGGAGCTGGAGAGC 360  
 Qy 361 GCAGCTCAACAGAGCAATAGGGAGATGGAGTTCACTGTTGACGAGATGGCTC 420  
 Db 361 GCAGCTCAACAGAGCAATAGGGAGATGGAGTTCACTGTTGACGAGATGGCTC 420  
 Qy 421 GATCCCTGACCTGTGATOCGGCCGCTGGCCTTCCAAGTGCCAGATTACGGAT 480  
 Db 421 GATCCCTGACCTGTGATOCGGCCGCTGGCCTTCCAAGTGCCAGATTACGGAT 480  
 Qy 481 GTGCATTGTAAGCAACTTGGGCAACTATGCTGAGAATGTACCCAGA 540  
 Db 481 GTGCATTGTAAGCAACTTGGGCAACTATGCTGAGAATGTACCCAGA 540  
 Qy 541 TGTATCATATCCTGTGGCTGAGAGGGFCCTTCAAGTATTCAGTCATCTTC 600  
 Db 541 TGTATCATATCCTGTGGCTGAGAGGGFCCTTCAAGTATTCAGTCATCTTC 600  
 Qy 601 TGCTTTGTCAGAAACACATGACCAAGCTGGCTCTGAAGATGTTACTACCGATAGC 660  
 Db 601 TGCTTTGTCAGAAACACATGACCAAGCTGGCTCTGAAGATGTTACTACCGATAGC 660  
 Qy 661 TTAAACTCAACCAATGTTTACTGAAATAACAAATGTTGTAATTCCTGAGTT 720  
 Db 661 TTAAACTCAACCAATGTTTACTGAAATAACAAATGTTGTAATTCCTGAGTT 720  
 Qy 721 TATTCTACTGTATAAAGGTTAATAATCATATAATCATATAATCTGGGATCATGGC 780  
 Db 721 TATTCTACTGTATAAAGGTTAATAATCATATAATCATATAATCTGGGATCATGGC 780  
 Qy 781 AGAGTTGTTGGGAGGGAAATGTTCACTGGTTCACTGAAATAATCCAAAAGTT 840  
 Db 781 AGAGTTGTTGGGAGGGAAATGTTCACTGGTTCACTGAAATAATCCAAAAGTT 840  
 Qy 841 ATTTCCTCAGAAAATCAATAAAGTTGCACTGTTTATCTTAAACATTTTAAAA 900  
 Db 841 ATTTCCTCAGAAAATCAATAAAGTTGCACTGTTTATCTTAAACATTTTAAAA 900  
 Qy 901 CCACCTGTTGAAATGGGACTGGCAATTTGCACTATACTATAATATAATTTAA 960

RESULT 2  
 AR167384 LOCUS AR167384 Sequence 20 from patent US 6287569.  
 DEFINITION Sequence 20 from patent US 6287569.  
 VERSION AR167384.1 GI:17903161  
 KEYWORDS Unknown  
 ORGANISM Unclassified  
 REFERENCE (bases 1 to 1004)  
 AUTHORS Kipps, T.J. and Wu, Y.  
 TITLE Vaccines with enhanced intracellular processing  
 JOURNAL US 6287569-A 20 11-SEP-2001;  
 FEATURES Location/Qualifiers  
 source 1. 1004  
 /organism="unknown"  
 /mol\_type="unassigned DNA"

ORIGIN

Query Match Score 100.0%; Score 1004; DB 6; Length 1004;  
 Best Local Similarity 100.0%; Prod. No. 3.7e-223;  
 Matches 1004; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCCCAATTAGGGTTCGGTATCCTCGTAACTGTTAGAGAAC 60  
 Db 1 CGCCCAATTAGGGTTCGGTATCCTCGTAACTGTTAGAGAAC 60  
 Qy 61 AGGAGAGGGGAGCTGGGGCTGGCTGAGCTGCTGAGCTGAGGAGCC 60  
 Db 61 AGGAGAGGGGAGCTGGGGCTGGCTGAGCTGCTGAGGAGCC 60  
 Qy 121 GTGGTGGCACACAGATGCCAGCTGGAGCTGGAGCTGGAGCTGAGG 120  
 Db 121 GTGGTGGCACACAGATGCCAGCTGGAGCTGGAGCTGGAGCTGAGG 120  
 Qy 181 AGTGGGGCGAGCTGGAGCTGAGATGCCGGCAAGGGCTTTCGGCATTTGTGCCAGC 240  
 Db 181 AGTGGGGCGAGCTGGAGCTGAGATGCCGGCAAGGGCTTTCGGATTTGTGCCAGC 240  
 Qy 241 TGCCTCAAGCCAGGTGATGAAGAGGACTCCCTGTTGACTGGTGGAGCTG 300  
 Db 241 TGCCTCAAGCCAGGTGATGAAGAGGACTCCCTGTTGACTGGTGGAGCTG 300  
 Qy 301 AAGACGGCACAGCTGTTGACTCTTGAGCTGGAGCTGGAGAGC 360  
 Db 301 AAGACGGCACAGCTGTTGACTCTTGAGCTGGAGCTGGAGAGC 360  
 Qy 361 GCAGCTCAACAGAGCAATAGGGAGATGGAGTTCACTGTTGACGAGATGGCTC 420  
 Db 361 GCAGCTCAACAGAGCAATAGGGAGATGGAGTTCACTGTTGACGAGATGGCTC 420  
 Qy 421 GATCCCTGACCTGTGATOCGGCCGCTGGCCTTCCAAGTGCCAGATTACGGAT 480  
 Db 421 GATCCCTGACCTGTGATOCGGCCGCTGGCCTTCCAAGTGCCAGATTACGGAT 480  
 Qy 481 GTGCATTGTCAGAAACACATGACCAAGCTGGCTCTGAAGATGTTACTACCGATAGC 540  
 Db 481 GTGCATTGTCAGAAACACATGACCAAGCTGGCTCTGAAGATGTTACTACCGATAGC 540  
 Qy 541 TGTATCATATCCTGTGGCTGAGAGGGFCCTTCAAGTATTCAGTCATCTTC 600  
 Db 541 TGTATCATATCCTGTGGCTGAGAGGGFCCTTCAAGTATTCAGTCATCTTC 600  
 Qy 601 TGCTTTGTCAGAAACACATGACCAAGCTGGCTCTGAAGATGTTACTACCGATAGC 660  
 Db 601 TGCTTTGTCAGAAACACATGACCAAGCTGGCTCTGAAGATGTTACTACCGATAGC 660  
 Qy 661 TTAAACTCAACCAATGTTTACTGAAATAACAAATGTTGTAATTCCTGAGTT 720  
 Db 661 TTAAACTCAACCAATGTTTACTGAAATAACAAATGTTGTAATTCCTGAGTT 720  
 Qy 721 TATTCTACTGTATAAAGGTTAATAATCATATAATCATATAATCTGGGATCATGGC 780  
 Db 721 TATTCTACTGTATAAAGGTTAATAATCATATAATCATATAATCTGGGATCATGGC 780  
 Qy 781 AGAGTTGTTGGGAGGGAAATGTTCACTGGTTCACTGAAATAATCCAAAAGTT 840  
 Db 781 AGAGTTGTTGGGAGGGAAATGTTCACTGGTTCACTGAAATAATCCAAAAGTT 840  
 Qy 841 ATTTCCTCAGAAAATCAATAAAGTTGCACTGTTTATCTTAAACATTTTAAAA 900  
 Db 841 ATTTCCTCAGAAAATCAATAAAGTTGCACTGTTTATCTTAAACATTTTAAAA 900  
 Qy 901 CCACCTGTTGAAATGGGACTGGCAATTTGCACTATACTATAATATAATTTAA 960

RESULT 2  
 AR167384 LOCUS AR167384 Sequence 20 from patent US 6287569.  
 DEFINITION Sequence 20 from patent US 6287569.  
 VERSION AR167384.1 GI:17903161  
 KEYWORDS Unknown  
 ORGANISM Unclassified  
 REFERENCE (bases 1 to 1004)  
 AUTHORS Kipps, T.J. and Wu, Y.  
 TITLE Vaccines with enhanced intracellular processing  
 JOURNAL US 6287569-A 20 11-SEP-2001;  
 FEATURES Location/Qualifiers  
 source 1. 1004  
 /organism="unknown"  
 /mol\_type="unassigned DNA"

ORIGIN

Query Match Score 100.0%; Score 1004; DB 6; Length 1004;  
 Best Local Similarity 100.0%; Prod. No. 3.7e-223;  
 Matches 1004; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCCCAATTAGGGTTCGGTATCCTCGTAACTGTTAGAGAAC 60  
 Db 1 CGCCCAATTAGGGTTCGGTATCCTCGTAACTGTTAGAGAAC 60  
 Qy 61 AGGAGAGGGGAGCTGGGGCTGGCTGAGCTGCTGAGGAGCC 60  
 Db 61 AGGAGAGGGGAGCTGGGGCTGGCTGAGCTGCTGAGGAGCC 60  
 Qy 121 GTGGTGGCACACAGATGCCAGCTGGAGCTGGAGCTGGAGCTGAGG 120  
 Db 121 GTGGTGGCACACAGATGCCAGCTGGAGCTGGAGCTGGAGCTGAGG 120  
 Qy 181 AGTGGGGCGAGCTGGAGCTGAGATGCCGGCAAGGGCTTTCGGCATTTGTGCCAGC 240  
 Db 181 AGTGGGGCGAGCTGGAGCTGAGATGCCGGCAAGGGCTTTCGGATTTGTGCCAGC 240  
 Qy 241 TGCCTCAAGCCAGGTGATGAAGAGGACTCCCTGTTGACTGGTGGAGCTG 300  
 Db 241 TGCCTCAAGCCAGGTGATGAAGAGGACTCCCTGTTGACTGGTGGAGCTG 300  
 Qy 301 AAGACGGCACAGCTGTTGACTCTTGAGCTGGAGCTGGAGAGC 360  
 Db 301 AAGACGGCACAGCTGTTGACTCTTGAGCTGGAGCTGGAGAGC 360  
 Qy 361 GCAGCTCAACAGAGCAATAGGGAGATGGAGTTCACTGTTGACGAGATGGCTC 420  
 Db 361 GCAGCTCAACAGAGCAATAGGGAGATGGAGTTCACTGTTGACGAGATGGCTC 420  
 Qy 421 GATCTCTGACCTGTGATGGAGCTGGCTCTGAAGGGACTCCCTGTTGACTGGTGGAGCTG 480  
 Db 421 GATCTCTGACCTGTGATGGAGCTGGCTCTGAAGGGACTCCCTGTTGACTGGTGGAGCTG 480  
 Qy 481 GTGCATTGTCAGAAACACATGACCAAGCTGGCTCTGAAGATGTTACTACCGATAGC 540  
 Db 481 GTGCATTGTCAGAAACACATGACCAAGCTGGCTCTGAAGATGTTACTACCGATAGC 540  
 Qy 541 TGTATCATATCCTGTGGCTGAGAGGGFCCTTCAAGTATTCAGTCATCTTC 600  
 Db 541 TGTATCATATCCTGTGGCTGAGAGGGFCCTTCAAGTATTCAGTCATCTTC 600  
 Qy 601 TGCTTTGTCAGAAACACATGACCAAGCTGGCTCTGAAGATGTTACTACCGATAGC 660  
 Db 601 TGCTTTGTCAGAAACACATGACCAAGCTGGCTCTGAAGATGTTACTACCGATAGC 660  
 Qy 661 TTAAACTCAACCAATGTTTACTGAAATAACAAATGTTGTAATTCCTGAGTT 720  
 Db 661 TTAAACTCAACCAATGTTTACTGAAATAACAAATGTTGTAATTCCTGAGTT 720  
 Qy 721 TATTCTACTGTATAAAGGTTAATAATCATATAATCATATAATCTGGGATCATGGC 780  
 Db 721 TATTCTACTGTATAAAGGTTAATAATCATATAATCATATAATCTGGGATCATGGC 780  
 Qy 781 AGAGTTGTTGGGAGGGAAATGTTCACTGGTTCACTGAAATAATCCAAAAGTT 840  
 Db 781 AGAGTTGTTGGGAGGGAAATGTTCACTGGTTCACTGAAATAATCCAAAAGTT 840  
 Qy 841 ATTTCCTCAGAAAATCAATAAAGTTGCACTGTTTATCTTAAACATTTTAAAA 900  
 Db 841 ATTTCCTCAGAAAATCAATAAAGTTGCACTGTTTATCTTAAACATTTTAAAA 900  
 Qy 901 CCACCTGTTGAAATGGGACTGGCAATTTGCACTATACTATAATATAATTTAA 960





541 TGATATCAATTATCCTTGTGAGGGCGCTTCAGGTTAGTCATCTTCC 600  
 601 TGCCTTGTCAGAACACATTGACCAAGCTCCTGAAAGATGTAAGTTTACTAGC 660  
 601 TGCCTTGTCAGAACACATTGACCAAGCTCCTGAAAGATGTAAGTTTACTAGC 660  
 Db 661 TTAAACTTCAACAACTTACTGAAATAAACAAATGTTAAATTCCCTGAGT 720  
 QY 661 TTAAACTTCAACAACTTACTGAAATAAACAAATGTTAAATTCCCTGAGT 720  
 Db 721 TAATCTACTGTTAAAGGTTAAACATATCATTAATCTAGGATTCTGCC 780  
 QY 721 TAATCTACTGTTAAAGGTTAAACATATCATTAATCTAGGATTCTGCC 780  
 Db 721 TAATCTACTGTTAAAGGTTAAACATATCATTAATCTGGGGATCATGCC 780  
 QY 781 AGAGATTGTTGGGGAGGAAGTTAAACGGTTTCACTGAAATTAAATCAA 840  
 Db 781 AGAGATTGTTGGGGAGGAAGTTAAACGGTTTCACTGAAATTAAATCAA 840  
 QY 841 ATTTCCTAGAAAAATCAATAAAGTTGATGTTTCACTGTTAAACATTTTAAA 900  
 Db 841 ATTTCCTAGAAAAATCAATAAAGTTGATGTTTCACTGTTAAACATTTTAAA 900  
 QY 901 CCACCTGAAATGTTAAATGGGACTGAAAGTTAAACGGTTTCACTGAAATTAAATTA 960  
 Db 901 CCACCTGAAATGTTAAATGGGACTGAAAGTTAAACGGTTTCACTGAAATTAAATTA 960  
 QY 961 TTAAAAAGTCATACTGTTACATTTACACTAAAGGC 1004  
 Db 961 TTAAAAAGTCATACTGTTACATTTACACTAAAGGC 1004  
 RESULT 6  
 HSU19180 HSU19180 1004 bp mRNA linear PRI 13-APR-1.995  
 LOCUS Human B melanoma antigen (BAGE) mRNA, complete cds.  
 DEFINITION Human B melanoma antigen (BAGE) mRNA, complete cds.  
 VERSION U19180  
 VERSION U19180..1 GI:726039  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 1004)  
 AUTHORS Boel,P., Wildmann,C., Sensi,M.L., Brassier,R., Renauld,J.C.,  
 Coulier,P., Boon,T. and van der Bruggen,P.  
 TITLE BAGE: a new gene encoding an antigen recognized on human melanomas  
 JOURNAL Immunity 2 (2), 167-175 (1995)  
 MEDLINE 95202592  
 PUBMED 7895173  
 REFERENCE 2 (bases 1 to 1004)  
 AUTHORS Boel,P.  
 TITLE Direct Submission  
 JOURNAL Submitted (28-DEC-1994) Pascale Boel, Ludwig Institute for Cancer  
 Research, Avenue Hippocrate 74, Brussels, B-1200, Belgium  
 FEATURES Location/Qualifiers  
 source 1..1004  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /isolate="patient MZ2"  
 /db\_xref="patient: 9606"  
 /sex="Female"  
 /cell\_line="MZ2-MEL-43"  
 /tissue\_type="melanoma"  
 /dev\_stage="adult"  
 gene 1..1004  
 /gene="BAGE"  
 CDS 201..332  
 /gene="BAGE"  
 /standard\_name="B melanoma antigen"  
 /codon\_start=1  
 /product="B melanoma antigen"  
 /protein\_id="AAC50123.1"  
 /protein\_id="AAC50123.1"

/db\_xref="GI:7260400"  
 /translation="MAARAVFLALSAOLJQARLMKEESPVWSWRLEPDGDIALCFI"  
 repeat\_region 100.0% Score 1004; DB 9; Length 1004;  
 Best Local Similarity 100.0%; Pred. No. 3..7..223;  
 Matches 1004; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN

Query Match 1 CGCCATTATGGGTCTCCGGATATCTCCCGCTGAGCTTAGGGACC 60  
 Local Similarity 1 CGCCATTATGGGTCTCCGGATATCTCCCGCTGAGCTTAGGGACC 60  
 Matches 1 AGAGAAGGGGGAGCTGGAGGTGGAGCTGAGCTGCTCACTTGGGAG 120  
 保守性 61 AGAGAAGGGGGAGCTGGAGGTGGAGCTGAGCTGCTCACTTGGGAG 120  
 差异 61 AGAGAAGGGGGAGCTGGAGGTGGAGCTGAGCTGCTCACTTGGGAG 120  
 缺失 121 GTGGTGGCAAAGAGATGGGAGGTGGAGTGTAGGTTAGGGGGCTGAGGTAGG 180  
 插入 121 GTGGTGGCAAAGAGATGGGAGGTGGAGTGTAGGTTAGGGGGCTGAGGTAGG 180  
 QY 181 AGTGGGGCTGAGCTAAATGGGGCTGAGCTAAATGGGGCTGAGCTGGCCAGC 240  
 Db 181 AGTGGGGCTGAGCTAAATGGGGCTGAGCTAAATGGGGCTGAGCTGGCCAGC 240  
 QY 241 TGCTCCAAGCCAGGGCTGATGAAAGGGAGTCCCCTGCTGCTGAGCTGGAGCCNG 300  
 Db 241 TGCTCCAAGCCAGGGCTGATGAAAGGGAGTCCCCTGCTGCTGAGCTGGAGCCNG 300  
 QY 301 AAGACGGCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGAGCTGGAGCAGC 360  
 Db 301 AAGACGGCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGAGCTGGAGCAGC 360  
 QY 421 GATCTCTGACCTCGTATGGCCCTTCACTCTGAGGTGTCGAGCCACGGTGTGCTC 420  
 Db 421 GATCTCTGACCTCGTATGGCCCTTCACTCTGAGGTGTCGAGCCACGGTGTGCTC 420  
 QY 481 GTGCAATTGTAAGGACTTTGGACCACTATCAATGCTGCTGAGAGAAATGTACCCGA 540  
 Db 481 GTGCAATTGTAAGGACTTTGGACCACTATCAATGCTGCTGAGAGAAATGTACCCGA 540  
 QY 541 TGTATCATATTACCTTATGCTGAGGCCGCTCTTCAAGATTCACTCTTC 600  
 Db 541 TGTATCATATTACCTTATGCTGAGGCCGCTCTTCAAGATTCACTCTTC 600  
 QY 601 TGCTGTCCAGAACACATTGACCAAGCTCTGAAGATGTAAGTTTACTACGGTAGAC 660  
 Db 601 TGCTGTCCAGAACACATTGACCAAGCTCTGAAGATGTAAGTTTACTACGGTAGAC 660  
 QY 721 TATTACTCTGATTAAAGGTAATTAATCATATAATCTGAGGATCTCTGCC 780  
 Db 721 TATTACTCTGATTAAAGGTAATTAATCATATAATCTGAGGATCTCTGCC 780  
 QY 781 AGAGATGTTGGGGAGGAATGTTCAAGGTTCACTGAAATTTAAATCCTAAAGTT 840  
 Db 781 AGAGATGTTGGGGAGGAATGTTCAAGGTTCACTGAAATTTAAATCCTAAAGTT 840  
 QY 841 ATTCCTCAGAAATCAATTAGTTGATGTTTCACTGAAATTTAAATCCTAAAGTT 900  
 Db 841 ATTCCTCAGAAATCAATTAGTTGATGTTTCACTGAAATTTAAATCCTAAAGTT 900  
 QY 901 CCACGTAGATGTTAAATGGGACTGTCAGATTGCTGAGCTATATAAAATA 960  
 Db 901 CCACGTAGATGTTAAATGGGACTGTCAGATTGCTGAGCTATATAAAATA 960





DD	961	AGTATTCGACATATCATAATTAAAGTCATCGTATCACTTTA 1020	
Qy	993	CACTAAAAGCC 1004	
Db	1021	CACTAAAAGCC 1032	
	RESULT 1.0		
LOCUS	AF527550	2154 bp mRNA linear PRI 31-JAN-2003	
DEFINITION	Homo sapiens B melanoma antigen variant b (BAGE1) mRNA, Complete		
ACCESSION	AF527550		
VERSION	AF527550.1	GI:28173029	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 2154)		
AUTHORS	Ruault, M., Van der Bruggen, P., Brun, M.-E., Boyle, S., Roizes, G. and Sario, A.D.		
TITLE	New BAGE (B melanoma antigen) genes mapping to the juxtacentromeric regions of human chromosomes 13 and 21 have a cancer/testis expression profile		
JOURNAL	Eur. J. Hum. Genet. 10 (12), 833-840 (2002)		
MEDLINE	22349465		
PUBMED	12461691		
REFFERENCES	2 (bases 1 to 2154)		
AUTHORS	Ruault, M., Ventura, M., Galtier, N., Brun, M.-E., Archidiacono, N., Roizes, G. and De Sario, A.		
TITLE	BAGE, a gene family generated by juxtacentromeric reshuffling in the Hominidae lineage, is under selective pressure		
JOURNAL	Genomics (2003) In press		
REFERENCE	3 (bases 1 to 2154)		
AUTHORS	Ruault, M., van der Bruggen, P., Brun, M.-E., Boyle, S., Roizes, G. and De Sario, A.		
TITLE	Direct Submission		
JOURNAL	Submitted (08 JUL-2002) Institut de Génétique Humaine, CNRS UPR 1142, 141, rue de la Cardinale, Montpellier 34396, France		
FEATURES	Location/Qualifiers		
Source	1. .2154 /gene="BAGE1"		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="mRNA"		
	/chromosome="13"		
gene	1. .2154		
CDS	201. .332 /gene="BAGE1"		
	/note="rumor antigen; alternatively spliced"		
	/codon_start=1		
	/product="B melanoma antigen variant b"		
	/protein_id="PAO32670_1"		
	/db_xref="GI:28173030"		
	/translation="MAARAVFLAUSAQLIQRALMKEESPVSVRLEPVEDGTALCFIF"		
ORIGIN			
	Query Match 82.6%; Score 829.2; DB 9; Length 2154;		
	Best Local Similarity 92.1%; Pred. No. 2e-182; Mismatches 8; Indels 71; Gaps 2;		
	Matches 925; Conservative 0; Mismatches 8; Definition Homo sapiens B melanoma antigen variant e (BAGE1) mRNA, complete.		
Qy	1 CGCCAAATTAGGTCTCGGTATCTCCGGTGAAGCTTGTCGTTCCGGTTAGGGACC 60	DEFINITION	
Db	1 CGCCAAATTAGGTCTCGGTATCTCCGGTGAAGCTTGTCGTTCCGGTTAGGGACC 60	ACCESSION AF527553	
	1 AGGAGAAGGGAGCTGGAGCTGGCAACCGTGCTGTTAGGGACC 60	VERSION 1 GI:28173035	
Qy	1 AGGAGAAGGGAGCTGGCAACCGTGCTGTTAGGGACC 60	ORGANISM Homo sapiens (human)	
Db	1 AGGAGAAGGGAGCTGGCAACCGTGCTGTTAGGGACC 60	SOURCE AF527553	
	1 (bases 1 to 1068)	ORGANISM Homo sapiens	
	1 (bases 1 to 1068)	DEFINITION Homo sapiens B melanoma antigen variant e (BAGE1) mRNA, complete.	
	1 (bases 1 to 1068)	ACCESSION AF527553	
	1 (bases 1 to 1068)	VERSION 1 GI:28173035	
	1 (bases 1 to 1068)	REFERENCE 1	
	1 (bases 1 to 1068)	REFERENCES 1	



Assembly program: Phrap; version 0.960731  
 Consensus Quality: 157343 bases at least Q40  
 Consensus Quality: 166322 bases at least Q30  
 Consensus Quality: 171194 bases at least Q20  
 Insert size: 15000; agarose-fp  
 Insert size: 17496; sum-of-contigs  
 Quality coverage: 3.5 in Q20 bases; agarose-fp  
 Quality coverage: 3.7 in Q20 bases; sum-of-contigs

\* NOTE: This is a 'working draft' sequence. It currently consists of 38 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

	FEATURES	SOURCE
1	1313: contig of 1313 bp in length 1413: gap of 100 bp 1414: contig of 1269 bp in length 2682: gap of 100 bp 2782: gap of 100 bp 4276: contig of 1494 bp in length 4377: gap of 100 bp 5635: contig of 1259 bp in length 5636: 5735: gap of 100 bp 5736: 7093: contig of 1358 bp in length 7094: 7193: gap of 100 bp 7194: 9078: contig of 1885 bp in length 9079: 9178: gap of 100 bp 9179: 10415: contig of 1237 bp in length 10416: 10515: gap of 100 bp 10516: 12355: contig of 1840 bp in length 12356: 12455: gap of 100 bp 12456: 15119: contig of 2664 bp in length 15120: 15220: gap of 100 bp 18382: 18482: contig of 3163 bp in length 18483: 19774: gap of 100 bp 19775: 19874: contig of 1292 bp in length 19875: 20901: gap of 100 bp 20902: 21001: contig of 1027 bp in length 21002: 22663: gap of 100 bp 22664: 22763: contig of 1662 bp in length 22764: 24927: contig of 1292 bp in length 24928: 25028: gap of 100 bp 25028: 27690: contig of 2663 bp in length 27691: 27791: gap of 100 bp 27791: 30074: contig of 2284 bp in length 30075: 30175: gap of 100 bp 30175: 33195: contig of 2021 bp in length 33196: 33295: gap of 100 bp 33296: 34897: contig of 1602 bp in length 34898: 38603: contig of 3606 bp in length 38604: 38703: gap of 100 bp 38704: 42337: contig of 2164 bp in length 42438: 42538: gap of 100 bp 42538: 46993: contig of 4456 bp in length 46994: 47093: gap of 100 bp 47094: 51119: contig of 4526 bp in length 51120: 51219: gap of 100 bp 51219: 55559: contig of 3640 bp in length 55360: 55559: gap of 100 bp 55460: 60487: contig of 5028 bp in length 60488: 60587: gap of 100 bp 60588: 64049: contig of 4022 bp in length 64610: 64709: gap of 100 bp 64710: 69559: contig of 4850 bp in length 69550: 69559: gap of 100 bp 69660: 74140: contig of 4481 bp in length 74141: 74240: gap of 100 bp 74241: 78594: contig of 4354 bp in length 78595: gap of 100 bp	* 78695: 83792: contig of 5098 bp in length * 83793: 83892: gap of 100 bp * 83893: 84946: contig of 5604 bp in length * 84947: 85596: gap of 100 bp * 85597: 92298: contig of 7702 bp in length * 92299: 97299: gap of 100 bp * 97299: 103662: contig of 8264 bp in length * 103663: 105762: gap of 100 bp * 105763: 114041: contig of 8279 bp in length * 114042: 114141: gap of 100 bp * 114142: 122299: contig of 7958 bp in length * 122299: 130991: gap of 100 bp * 130992: 131091: contig of 8792 bp in length * 131092: 143872: contig of 12781 bp in length * 143873: 143972: gap of 100 bp * 143973: 155554: contig of 11882 bp in length * 155554: gap of 100 bp * 155555: 177896: contig of 21942 bp in length.  Location Qualifiers 1. 177896 /clone="RP11-63BN24" /mol_type="genomic DNA" /db_xref="taxon:9606" /map="4" /clone_lib="RPCI-11_Human_Male_BAC"  misc_feature 1. .1313 /note="assembly_fragment" 1414. 2682 /note="assembly_fragment" 2783. 4276 /note="assembly_fragment" 4377. 5635 /note="assembly_fragment" 5736. 7093 /note="assembly_fragment" 7194. 9078 /note="assembly_fragment" 9179. 10415 /note="assembly_fragment" 10416. 12355 /note="assembly_fragment" 12356. 15119 /note="assembly_fragment" 15120. 18382 /note="assembly_fragment" 18483. 19774 /note="assembly_fragment" clone_end:T7 vector_side:right"  misc_feature 19815. 20901 /note="assembly_fragment" 21002. 22663 /note="assembly_fragment" 22764. 24927 /note="assembly_fragment" 25028. 25028 /note="assembly_fragment" 27791. 30074 /note="assembly_fragment" 30175. 33195 /note="assembly_fragment" 33296. 34897 /note="assembly_fragment" 34898 /note="assembly_fragment" 38603 /note="assembly_fragment" 38604 /note="assembly_fragment" 38704 /note="assembly_fragment" 42337 /note="assembly_fragment" 42438 /note="assembly_fragment" 42538 /note="assembly_fragment" 46993 /note="assembly_fragment" 46994 /note="assembly_fragment" 47093 /note="assembly_fragment" 51119 /note="assembly_fragment" 51219 /note="assembly_fragment" 55559 /note="assembly_fragment" 55360 /note="assembly_fragment" 55559 /note="assembly_fragment" 60487 /note="assembly_fragment" 60488 /note="assembly_fragment" 64049 /note="assembly_fragment" 64610 /note="assembly_fragment" 64709 /note="assembly_fragment" 69559 /note="assembly_fragment" 69550 /note="assembly_fragment" 74140 /note="assembly_fragment" 74240 /note="assembly_fragment" 78594 /note="assembly_fragment" 78595 /note="assembly_fragment"

Query Match	53.0%	Score 532.4;	DB 2;	Length 177896;
Best Local Similarity	99.8%;	Pred. No. 3.2e-113;		
Matches	533;	Conservative	0;	Mismatches 1; Indels 0; Gaps 0;
LOCUS	AC133563	139063 bp	DNA	linear HTG 14-SEP-2002
DEFINITION	Homo sapiens chromosome 16 clone RP11-621B, WORKING DRAFT SEQUENCE, 15 unordered pieces.			
ACCESSION	AC133563			
VERSION	AC133563.1	GI:122867559		
KEYWORDS	HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.			
ORGANISM	Homo sapiens			
COMMENT	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1 (bases 1 to 139063)			
AUTHORS	DOE Joint Genome Institute.			
JOURNAL	Sequencing of Human Chromosome 16			
REFERENCE	Unpublished			
AUTHORS	2 (bases 1 to 139063)			
JOURNAL	DOE Joint Genome Institute.			
COMMENT	Submitted (14-SEP-2002) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA			
COMMENT	Center: Joint Genome Institute			
COMMENT	Center Code: JGI			
COMMENT	Web site: <a href="http://www.jgi.doe.gov">http://www.jgi.doe.gov</a>			
FEATURES	1			RESULT 13
Source	AC133563	139063 bp	DNA	linear HTG 14-SEP-2002
Location/Qualifiers	1 . 139063	Score 529.2;	DB 2;	Length 139063;
	1 . 139063	Pred. No. 1.7e-112;		
	1 . 139063	Mismatches 0;	Indels 0;	Gaps 0
ORIGIN				
Query Match	52.7%;	Score 529.2;	DB 2;	Length 139063;
Best Local Similarity	99.4%;	Pred. No. 1.7e-112;		
Matches	531;	Conservative	0;	Mismatches 3;
Qy	471	TTACAGATGTGATTGTAAACACTTGGACCACTATAATGCTGTAGAGAA	5303	
Db	86678	TTTAGCAGTGTGATTGTAAACACTTGGACCACTATAATGCTGTAGAGAA	86718	
Qy	531	TGTACCGAGATGATCATTCATTCTGTGTGTGAGGAGGGCTTCAGGATTCTAGT	590	
Db	86738	TGACCCAGATGATCATTCATTCTGTGTGTGAGGAGGGCTTCAGGATTCTAGT	86718	
Qy	591	CACATCTCTGCTTGTCCAGAACACATTGACAAAGCTCCTGAAAGATGTAAAGTTACT	650	
Db	86798	CACATCTCTGCTTGTCCAGAACACATTGACAAAGCTCCTGAAAGATGTAAAGTTACT	868	
Project Information				
Center Project Name:	430373			
Center clone name:	RPCI-11_621B			
Summary Statistics				

RESULT 14				FEATURES		Location/Qualifiers	
Qy	AC145613	source					
Qy	7711	CCCTGAGTGTATTCTACTGTATTAAAGGTATAATACTTAAATCTGAGG	770	*	4198	6161: contig of 1964 bp in length	
Db	86918	CCCTGAGTGTATTCTACTGTATTAAAGGTATAATACTTAAATCTGAGG	86977	*	6162	6262: contig of 5133 bp in length	
Qy	7711	GATCATTGCCAGAGATTGTTGGGGAAATGTTCAACGGTTTCATGAAATTAAAT	830	*	11394	11395: gap of unknown length	
Db	86978	GATCATTGCCAGAGATTGTTGGGGAAATGTTCAACGGTTTCATGAAATTAAAT	87037	*	11495	11494: gap of unknown length	
Qy	8311	CCAAAAAAGCTTATTCTCGAGAAAATCAAAATAAAGTTGTCATTTTCTTAAAC	890	*	13311	13312: gap of unknown length	
Db	87038	CCAAAAAAGCTTATTCTCGAGAAAATCAAAATAAAGTTGTCATTTTCTTAAAC	87097	*	13412	16041: contig of 2630 bp in length	
Qy	8911	ATTTAAACCACTGTGAAATGATGTAATAAGGACTGTGCGATTTCTGACATAC	950	*	16042	16141: gap of unknown length	
Db	87098	ATTTAAACCACTGTGAAATGATGTAATAAGGACTGTGCGATTTCTGACATAC	87157	*	16142	16488: contig of 10347 bp in length	
Qy	9511	TATAAAATATTAAAGGCAATAGTATCAACATCTTACACTAAAGGC	1004	*	26589	26588: gap of unknown length	
Db	87158	TATAAAATATTAAAGGCAATAGTATCAACATCTTACACTAAAGGC	87211	*	47630	47729: gap of unknown length	
RESULT 14		source		FEATURES		Location/Qualifiers	
Qy	AC145613	DEFINITION Homo sapiens chromosome 178300 bp DNA Linear FTG 19-JUL-2003	PROGRESS **, 14 unordered pieces.	/organism="Homo sapiens"			
ACCESSION				/mol_type="genomic DNA"			
VERSION				/db_xref="taxon:9606"			
KEYWORDS				/chromosome="UNK"			
SOURCE				/clone="RP11-1266BH24"			
ORGANISM	Homo sapiens			1..1534			
Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;				/note="assembly_name:Contig13"			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				/note="genomic DNA"			
REFERENCE	AC145613..1	GI:32996952		/note="taxon:9606"			
KEYWORDS	HTGS PHASEI.			/note="clone=RP11-1266BH24"			
SOURCE	Homo sapiens (human)			1..1534			
ORGANISM	Homo sapiens			/note="assembly_name:Contig13"			
Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;				/note="genomic DNA"			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				/note="taxon:9606"			
REFERENCE	1 (bases 1 to 178300)			/note="chromosome="UNK"			
AUTHORS	Wilson, R.K.			/note="clone=RP11-1266BH24"			
SOURCE	The sequence of Homo sapiens clone			1..1534			
ORGANISM	Unpublished			/note="assembly_name:Contig14"			
Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;				/note="genomic DNA"			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				/note="taxon:9606"			
REFERENCE	2 (bases 1 to 178300)			/note="clone=RP11-1266BH24"			
AUTHORS	Wilson, R.K.			1..1534			
TITLE	Direct Submission			/note="assembly_name:Contig15"			
JOURNAL	Submitted (19-JUL-2003) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA			/note="assembly_name:Contig16"			
COMMENT	-----	-----	-----	1..1534			
Center: Washington University Genome Sequencing Center				/note="assembly_name:Contig17"			
Center code: WUGSC				1..1534..1..1534			
Web site: http://Genome.wustl.edu/gsc/index.shtml				/note="assembly_name:Contig18"			
-----	-----	-----	-----	1..1534..1..1534			
Center project name: H_NH1266H24				/note="assembly_name:Contig19"			
-----	-----	-----	-----	1..1534..1..1534			
Summary Statistics				1..1534..1..1534			
Sequencing vector: M13: 0%				1..1534..1..1534			
Chemistry: Dye-Primer ET; 0% of reads				1..1534..1..1534			
Assembly program: Phrap; version 0.990319				1..1534..1..1534			
Consensus quality: 171980 bases at least Q40				1..1534..1..1534			
Consensus quality: 173066 bases at least Q20				1..1534..1..1534			
-----	-----	-----	-----	1..1534..1..1534			
NOTE: This is a 'working draft' sequence. It currently consists of 14 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.				1..1534..1..1534			
* * consists of 14 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.				1..1534..1..1534			
* 1 1534: contig of 1534 bp in length				1..1534..1..1534			
* 1634: gap of unknown length				1..1534..1..1534			
* 1635 4097: contig of 2463 bp in length				1..1534..1..1534			
* 4098: gap of unknown length				1..1534..1..1534			
ORIGIN	-----	-----	-----	1..1534..1..1534			
Query Match	52.4%			1..1534..1..1534			
Best Local Similarity	99.1%			1..1534..1..1534			
Matches	529			1..1534..1..1534			
Conservative	0			1..1534..1..1534			
Indels	0			1..1534..1..1534			
Gaps	0			1..1534..1..1534			
-----	-----	-----	-----	1..1534..1..1534			
Qy	471	TTACACGATGAGTGTATATTCCCTGAGGACCTTGAGCTCATCATGCTGTGAGGAGCGCTCTTCAGGATTCTG	590				
Db	165850	TTTCACGATGATGCTTAACTGAGCTTGTGAGGAGCGCTCTTCAGGATTCTGAGGAGCGCTCTTCAGGATTCTG	165909				
Qy	531	TGTACCCAGATGATGCTTAACTGAGCTTGTGAGGAGCGCTCTTCAGGATTCTGAGGAGCGCTCTTCAGGATTCTG	165969				
Db	165910	TGTACCCAGATGATGCTTAACTGAGCTTGTGAGGAGCGCTCTTCAGGATTCTGAGGAGCGCTCTTCAGGATTCTG	165969				
Qy	591	CACATCTCTCTGCTTGTGAGGAGCGCTCTTCAGGATTCTGAGGAGCGCTCTTCAGGATTCTGAGGAGCGCTCTTCAGGATTCTG	650				

Db	165070	CACATCTCCGTCTGCGAAACATTGACCAAGCTCTGTAATTACTTACT 166029	Qy	531	TGTACCCAGATGTTATCATTATCCTTGCTGAGGAGGCCGGTCCTTCAGGATTCAGT 590
Qy	651	ACGATGACTTTAACCTCAACCAATGTTATTACTGAAATAACAAATGTTAATT 710	Db	77658	TGACCCAGATCTATTCTGCTTCAGATTCAGTTACT 77629
Db	166030	ACGATGACTTTAACCTCAACCAATGTTATTACTGAAATAACAAATGTTAATT 166089	Qy	591	CACATCTCCGTCTTCTGCTTCAGAACACATTGACCAAGCTCTGAAAGATGTTAAGTTACT 650
Qy	711	CCCTGAGGTATTCTACTGTATTAAAGGTAATAATCATTAATAATGTTGAGG 770	Db	77628	CACATCTCCGTCTTCTGCTTCAGAACACATTGACCAAGCTCTGAAAGATGTTAAGTTACT 77569
Db	166090	CCCTGAGGTATTCTACTGTATTAAAGGTAATAATCATTAATAATGTTGAGG 166149	Qy	651	ACCATGACTTTAACCTCAACCATGTTAAGTTACTGAAATAAAATGTTGAATT 710
Qy	771	GATCATGGAGATTGTTGGGAGGAAATGTTCAACGGTTTCATTGAAATTAAAT 830	Db	77568	ACGGATGACTTTAACCTCAACCATGTTAAGTTAACTGAAATAAAATGTTGAATT 77509
Db	166150	GATCGTTGCCAGAGTTGTTGGGGAGGAAATGTTCAACGGTTTCATTGAAATTAAAT 166209	Qy	711	CCCTGACTGTATTACTGTATTAAAGGTTAACTCATTAATACTGTTGAGG 770
Qy	831	CCAAAAGGTATTCTCCAGAAAATCAAAACTGTTTATTCATTAACAC 890	Db	77508	CCCTGAGGTATTACTGTATTAAAGGTTAACTCATTAATACTGTTGAGG 77449
Db	166210	CCAAAAGGTATTCTCCAGAAAATCAAAACTGTTTATTCATTAACAC 890	Qy	771	GATCATGGCCAGAGATTGTTGGGAGGAAATGTTATCAACGGTTTCATTGAAATTAAAT 830
Qy	891	ATTTTAAACCACTGTAGATGATGAAATAGGGACTGTCAGTTCTGCACATAC 950	Db	77448	GATCGTTGCCAGAGATTGTTGGGGAGAAATGTTATCAACGGTTTCATTGAAATTAAAT 77389
Db	166270	ATTTTAAACCACTGTAGAAAGGTTAACTGTTTACACTAAAGGCC 1004	Qy	831	CCAAAAGGTATTCTCCAGAAAATCAAAACTGTTTATTCATTAACAC 890
Qy	951	TATAAAATTATAAAAGGTAATACTGTTTACACTAAAGGCC 1044	Db	77388	CCAAAAGGTATTCTCCAGAAAATCAAAACTGTTTATTCATTAACAC 1004
Db	166330	TATAAAATTATAAAAGGTAATACTGTTTACACTAAAGGCC 166383	Qy	891	ATTTAAAAACCCTGTAGATGTTAACTGAGTTCTGACATATACTGACATATACT 950
			Db	77328	ATTTAAAAACCCTGTAGATGTTAACTGAGTTCTGACATATACTGACATATACT 77269
RESULT 15			Qy	951	TATAAAATTATAAAAGGTAATACTGTTTACACTAAAGGCC 1004
AF155875/C		AF155875 107701 bp DNA linear HTG 31-DEC-2001	Db	77268	TATAAAATTATAAAAGGTAATACTGTTTACACTAAAGGCC 77215
LOCUS		DEFINITION Homo sapiens clone MC169, *** SEQUENCING IN PROGRESS ***, 2 ordered pieces			
ACCESSION		AF155875			
VERSION		AF155875.1 GI:18000259			
KEYWORDS		HTG; HTGS; PHASE2.			
SOURCE		Homo sapiens (human)			
ORGANISM					
REFERENCE					
AUTHORS					
TITLE		Xia,J.-H., Huang,L., Tang,D.-S., Dai,H.-P., Pan,Q. and Long,Z.-G.			
JOURNAL		Direct Submission			
		Submitted (18-MAY-1999) National Lab of Medical Genetics of China,			
COMMENT		Hunan Medical University, 88 Xiang-ya Road, Changsha, Hunan 410078, P.R. China			
		* NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. Gaps between the contigs			
		* are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have			
		* provided by the submitter.			
		* This sequence will be replaced			
		* by the finished sequence as soon as it is available and			
		* the accession number will be preserved.			
		* * 105436: contig of 105436 bp in length			
		* * 105536: gap of unknown length			
		* * 105537 107701: contig of 2165 bp in length.			
FEATURES		Location/Qualifiers			
source		1..107701			
		/organism="Homo sapiens"			
		/mol_type="Genomic DNA"			
		/db_xref="Taxon:9606",			
		/clone="MC169"			
ORIGIN					
Query Match		52.2% Score 524.4; DB 2; Length 107701;			
		Best Local Similarity 98.9%; Pred. No. 2,3e-111;			
		Matches 526; Conservative 0; Mismatches 6; Indels 0; Gaps 0;			
Qy	471	TTACGGATGTGCAATTGTAAGGACTTGGACCTTGGACCTAAATGTTGAGAGAAA 530			
Db	77748	TTTCAGGAAATGTCATTGTAAGGACTTGGACCTTGGACCTAAATGTTGAGAGAAA 77689			

This Page Blank (uspto)

GenCore version 5.1.6  
(c) 1993 - 2004 Compugen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: July 8, 2004, 23:14:12 ; Search time 633 Seconds

(without alignments)  
6738.058 Million cell updates/sec

Title: US-10-081-108-1

Perfect score: 1004

Sequence: 1 CGCAATTAGGGTCTCGG... ATCTTTACACTAAAAAGCC 1004

Scoring table: IDENTIV\_NUC

GapOp 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters:

6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_29Jan04:\*

1: GeneseqN1980s:\*

2: GeneseqN1990s:\*

3: GeneseqN2000s:\*

4: GeneseqN2001as:\*

5: GeneseqN2001bs:\*

6: GeneseqN2002s:\*

7: GeneseqN2003as:\*

8: GeneseqN2003bs:\*

9: GeneseqN2003cs:\*

10: GeneseqN2004s:\*

## ALIGNMENTS

Adb72290 Human WNT

Aas46736 Tumour su

Ab166140 Lung canc

Abn93947 Gene #445

Abt44365 Partial g

Aas3104 Human Oes

Abz68064 Human sec

Adc20955 Human sec

Abn80533 Human sec

AbnB0353 Human P45

Abv1929 Human pro

Abv9004 Human pro

Aal03915 Human rep

Aal03913 Human rep

Aas40313 DNA encod

Aas40315 DNA encod

Aac12905 Human sec

Aak68680 Human imm

Ach2606 Human foie

Abv59611 Human pro

Abn6177 Human can

Aak78891 Human imm

## RESULT

1 AAT36382

standard; cdNA; 1004 BP.

XX AAT36382;

AC

XX

DT 04-DEC-1996

(first entry)

DE BAGE tumour rejection antigen precursor cDNA.

XX

KW BAGE; tumour rejection antigen precursor; TRAP; MHC;

major histocompatibility complex; HLA-Cw\*1601; melanoma; metastasis;

diagnosis; therapy; vaccine; ss.

XX Homo sapiens.

OS

XX

FH Key

FT primer\_bind

FT Location/Qualifiers

complement(100. .117)

/\*tag= ^

/product= 'primer VDB85'

201. .332

FT /\*tag= b

/product= "tumour rejection antigen precursor"

FT 204. .230

FT /\*tag= c

/product= "tumour rejection antigen"

FT 367. .385

FT /\*tag= d

/product= "tumour rejection antigen"

FT 385. .484

FT /\*tag= e

/note= 'Alu repeat'

XX WO9655511-A1.

XX PR 16-FEB-1995;

XX XX 95US-0039360.

(LUDWIG INST CANCER RES.

XX PA

XX XX

PI Boel P,

Boonfaelleur T,

Van Der Bruggen P,

Coulie P;

PI Renaud J;

XX

## SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No. Score Query Match Length DB ID Description

1 1004 100.0 1004 2 AAT36382 Rat36382 BAGE tumo

2 1004 100.0 1004 2 ABQ76214 Human tum

3 1004 100.0 1004 6 ABV72285 Nucleotid

4 1004 100.0 1004 9 ADD25519 Birding d

5 958 95.4 1032 2 AAQ81005 Hagg1005 BAGE tumo

6 52.4 52.2 11162 6 ABL55833 Human mut

7 52.4 52.2 11162 6 AB154504 Gene vect

8 52.4 52.2 11162 6 AAK93658 DNA of th

9 52.4 52.2 11162 6 ABA93937 Human sec

10 52.4 52.2 11162 6 AAB77500 Human TNK

11 52.4 52.2 107602 6 AAK9657 DNA of th

12 52.4 52.2 107612 6 ABL54503 Human PAC

13 181.8 18.1 57728 4 AAC87588 Human sp1

14 149.8 14.9 6292 4 AAS46735 Human tum

15 106.2 10.6 251 3 AAC19397 Human sec

16 8.7 130 4 AAL36471 Human mis

17 8.7 130 4 ABX59459 cDNA enco

18 8.7 2028 5 ABV2788 Human pro

19 8.7 21423 4 AAI36470 Human mus

20 8.7 21423 4 ABX59458 cDNA enco

21 8.7 476 6 ABV7857 Human col

22 8.6 72332 8 ADA02552 Human WNT

WPI; 1996-392411/39.	DR	DR	P-FSDB; AAW0152.
XX	XX	XX	Tumour rejection antigen precursor (TRAP) and gene - useful to develop prods. for diagnosis and treatment of disorders characterised by TRAP, partic. melanomas.
PPT	PPT	PPT	
XX	XX	XX	Claim 1; Page 27-28; 44pp; English.
CCC	CCC	CCC	A cDNA clone (AAT36382), designated CDNA-AD5, codes for the BAGE tumour rejection antigen precursor (TRAP) (AAW02152). MHC molecule HLA-Cw*1601 presents a tumour rejection antigen derived from the BAGE TRAP. The cDNA was identified by cotransferring HLA-Cw*1601 cDNA with a cDNA library derived from melanoma cell line MZ2-MEL43 into COS-7 cells and isolating clones capable of stimulating prodn. of tumour necrosis factor by cytotoxic T-lymphocytes. In 600 samples of tumours, the BAGE Gene was expressed mainly in melanomas (esp. in metastatic lesions), bladder carcinomas and mammary carcinomas. The isolated gene may be used diagnostically or for prodn. of the BAGE TRAP
CCS	CCS	CCS	Sequence 1004 BP; 273 A; 204 C; 255 G; 272 T; 0 U; 0 Other;
XX	XX	XX	SQ

Qy	721	TATTCTACTTGTATTAAAGGTTAATATACTATACTTAAATCATGGGCATCATGGCC	780
D <sub>b</sub>	721	TATTCTACTTGTATTAAAGGTTAATATACTATACTTAAATCATGGGCATCATGGCC	780
Qy	781	AGAGATTTGGGGAGGAATAGTTCAACCGTTCTATGTAATTAAATCaaaaAGTT	840
D <sub>b</sub>	781	AGAGATTTGGGGAGGAATAGTTCAACCGTTCTATGTAATTAAATCaaaaAGTT	840
Qy	841	ATTTCTCTAGAAAAATCAAATAAGTTGCATGTTTTATCTCTAAAACATTTTAAA	900
D <sub>b</sub>	841	ATTTCTCTAGAAAAATCAAATAAGTTGCATGTTTTATCTCTAAAACATTTTAAA	900
Qy	901	CCACTGTGAAGATGTAAATGGGATGTGGAGTATTCTGACATATACTTAAATTA	960
D <sub>b</sub>	901	CCACTGTGAAGATGTAAATGGGATGTGGAGTATTCTGACATATACTTAAATTA	960
Qy	961	TTAAAAGTCATCGATTCACTTACACTTTAACATAAAAGCC	1.004
D <sub>b</sub>	961	TTAAAAGTCATCGATTCACTTACACTTTAACATAAAAGCC	1.004

**RESULT 2**

ABQ76214	ID	ABQ76214	standard;	DNA;	1004	BP.									
XX															
AC		ABQ76214;													
XX	DT	21-OCT-2002	(first entry)												
XX	DE	Human tumour antigen BAGE DNA.													
XX	KW	Tumour antigen; human; vaccine; cellular immune response; immunogen;													
XX	KW	cancer; tumour; BAGE; ds.													
XX	OS	Homo sapiens.													
XX	PN	US6287569-B1.													
XX	PD	11-SEP-2001.													
XX	PP	06-APR-1998;													
XX	PR	98US-00056105.													
XX	PR	10-APR-1997;													
XX	PA	97US-004346/P.													
XX	PA	(REGC ) UNIV CALIFORNIA.													
XX	PI	Kipps TJ, Wu Y;													
XX	WPI	1998-583198/49.													
XX	DR	Generating cellular immune response in Patient to target protein -													
PT	comprises introducing vector with nucleotide sequence encoding immunogen														
PT	comprising protein processing signal into cell of Patient.														
XX	Disclosure	Co 1 73-76; 61PP; English.													
XX	CC	This invention describes a novel method for generating a cellular immune response in a patient to a target protein or its fragment. The method involves introducing a vector containing a nucleotide sequence encoding a chimeric immunogen comprising a protein processing signal and the target protein or its fragment. The immunogen is produced by the cells and processed so that the target protein or its fragment is presented to the patients immune system and a cellular immune response is initiated. The method and vectors can be used as a form of vaccination and could be used to generate a cellular immune response in patients to, e.g. cancerous tumours. The cellular immune response is the predominant immune response in the patient. This sequence represents a DNA fragment which encodes the human tumour antigen BAGE described in the method of the invention. Note: The information in this spec has been previously disclosed in WO199845414 however this spec contained no sequence information													
XX	Sequence	1004	BD	273	A	204	C	255	G	272	T	0	U	0	Other:

Query	Match	100.0%	Score 1004;	DB 2;	Length 1004;
	Best Local Similarity	100.0%;	Pred. No. 1-8e-250;	Mismatches 0;	Indels 0;
	Conservative	0;	Mismatches -0;		
Qy					
Db	1 CGCCAATTAGGGTCTCCGGATATCTCCCGTGAATGCTGACTCTGTGCCGGTTAA				
Qy	1 CSCCAATTAGGGTCTCCGGATATCTCCCGTGAATGCTGACTCTGTGCCGGTTAA				
Db	61 AGAGAGGGGGAGCTGGAGGTGGAGCTGAGGTTGGAGCTACTACACCGTGGCTSTA				
Qy	61 AGAGAGGGGGAGCTGGAGGTGGAGCTGAGGTTGGAGCTACTACACCGTGGCTSTA				
Db	121 GCTGGCCAACTGGAGTGGAGCACTAAGATGGCGGCCAGAGCGGTTTTCTGCATGTC				
Qy	121 GCTGGCCAACTGGAGTGGAGCACTAAGATGGCGGCCAGAGCGGTTTTCTGCATGTC				
Db	181 AGTGGGGCTGGAGCACTAAGATGGCGGCCAGAGCGGTTTTCTGCATGTC				
Qy	181 AGTGGGGCTGGAGCACTAAGATGGCGGCCAGAGCGGTTTTCTGCATGTC				
Db	181 AGTGGGGCTGGAGCACTAAGATGGCGGCCAGAGCGGTTTTCTGCATGTC				
Qy	241 TGCTCCAGCAGGCTCATGAGAGAGGAGTCCCTGTGAGCTGAGGAGTTGGCTGAG				
Db	241 TGCTCCAGCAGGCTCATGAGAGAGTCCCTGTGAGCTGAGGAGTTGGCTGAG				
Qy	301 AAGACGCACAGCTCTGGCTCATGAGGAGATGGAGTTCACTGTGTAAGCCAGGAA				
Db	301 AAGACGCACAGCTCTGGCTCATGAGGAGATGGAGTTCACTGTGTAAGCCAGGAA				
Qy	361 GGAGCTCAACAGGAGATAAGGAGGATGGAGTTCACTGTGTAAGCCAGGAA				
Db	361 GGAGCTCAACAGGAGATAAGGAGGATGGAGTTCACTGTGTAAGCCAGGAA				
Qy	421 GATCTCTGACTCTGTGATCGGCCGACCTTGCCCTTCAAAAGTCGCCGAGATTAC				
Db	421 GATCTCTGACTCTGTGATCGGCCGACCTTGCCCTTCAAAAGTCGCCGAGATTAC				
Qy	481 GTGCAATTGGTGGACCACTTGAGGCCCTATCAANTGGCTGAGAGAAAATGTA				
Db	481 GTGCAATTGGTGGACCACTTGAGGCCCTATCAANTGGCTGAGAGAAAATGTA				
Qy	541 TGTATCATTAATCTTGTGCTGAGGCCGCTCTTCAAGATTCTCACTACAC				
Db	541 TGTATCATTAATCTTGTGCTGAGGCCGCTCTTCAAGATTCTCACTACAC				
Qy	601 TGCTTGTCCGAACAACATGACAAAGCTCCTGAAAGATCTAAATTACTACGGC				
Db	601 TGCTTGTCCGAACAACATGACAAAGCTCCTGAAAGATCTAAATTACTACGGC				
Qy	661 TTAAACACTTCGACCAATGTATTACTGAAAATACAAAATGTTGTTAAATTCCC				
Db	661 TTAAACACTTCGACCAATGTATTACTGAAAATACAAAATGTTGTTAAATTCCC				
Qy	721 TATTCFACTGTATTAAGGTTAAATACATATCATTAATCTGAGGATCTGAGGATC				
Db	721 TATTCFACTGTATTAAGGTTAAATACATATCATTAATCTGAGGATCTGAGGATC				
Qy	781 AGAGATGTTGGGAGGAAATGTTATCACGGTTCATGAAATTAACTCCAA				
Db	781 AGAGATGTTGGGAGGAAATGTTATCACGGTTCATGAAATTAACTCCAA				
Qy	841 ATTTCCTCAGAAAATCAAAATCAAAATCAAAATGTTATCTGAGATAACATT				
Db	841 ATTTCCTCAGAAAATCAAAATCAAAATCAAAATGTTATCTGAGATAACATT				
Qy	901 CCACCTGAGATGATGTTAAATGAGCTGAGCTGAGATAACTATA				
Db	901 CCACCTGAGATGATGTTAAATGAGCTGAGCTGAGATAACTATA				
Qy	961 TTAAGAAGTCATCAGTATTCAAACTCTTTACACTAAAAAGGCC 1004				
Db	961 TTAAGAAGTCATCAGTATTCAAACTCTTTACACTAAAAAGGCC 1004				

RESULT 3  
V72285 ABV72285 Standard; DNA; 1004 BP.  
ABV72285;  
16-DEC-2002 (first entry)  
Nucleotide sequence of human BAGE.  
Human; IFTT-2; chronic myelogenous leukemia; IAGE-1; BAGE; DDB1; ETS2; PIASy; PIASx-alpha; PIASx-beta; DAK3; gene; ss.  
Homo sapiens.  
Key Location/Qualifiers  
CDS 201 . 332  
/\*tag= a  
/product = "BAGE"  
  
WO20020747-A1.  
12-SEP-2002.  
01-MAR-2002; 2002WO-JP001901.  
01-MAR-2001; 2001JP-00056438.  
(FUJI ) FUJISAWA PHARM CO LTD.  
Mano H;  
WPI; 2002-682911/73.  
P-PSDp; ABB7834.  
Measuring the expression profile of genes in a cell or tissue sample for diagnosis of chronic myelogenous leukemia and identification of agents for its treatment.  
Example 3; Page 38-39; 97pp; Japanese.  
The present sequence encodes human BAGE. The expression level of the gene is used in the method of the invention. The specification describes a method of examining chronic myelogenous leukemia. The method comprises measuring the expression level of a gene selected from IFTT-2, LAGE-1, BAGE, DDB1, ETS2, PIASy, PIASx alpha, PIASx-beta and DAK3, or determining the expression profile of a group of genes including one or more of these genes, in a cell or tissue sample from a chronic myelogenous leukemia patient. The method is used for the diagnosis, treatment and prevention of chronic myelogenous leukemia.  
Sequence 1004 BP; 273 A; 204 C; 255 G; 272 T; 0 U; 0 Other;  
Query Match 100.0%; Score 1004; DB 6; Length 1004;  
Best Local Similarity 100.0%; Pred. No. 1.8e-250;  
Matches 1004; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 CGCCAAATTAGGTCTCCGGTATCTCCGGTCTGAGCTGCTGTCCTGCTCACTGGATG 60  
1 CGCCAAATTAGGTCTCCGGTATCTCCGGTCTGAGCTGCTGTCCTGCTCACTGGATG 60  
1 AGGAGAAGGGGAGCTGGCTGAGCTGCTGAAACCGTGGCTGCTCACTGGATG 120  
1 AGGAGAAGGGGAGCTGGCTGAGCTGCTGAAACCGTGGCTGCTCACTGGATG 120  
121 GTGGTGGCAACAGAGATGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGG 180  
121 GTGGTGGCAACAGAGATGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGG 180  
181 AGTGGGGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGG 240  
181 AGTGGGGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGG 240



			Query	Match	Score	DB	Length
Db	601	TGCTTGCCAGAACATTTGACCAAGTCTGTAAAGTAACCTACAG 660	QY	601 TTTTAACTCACCATTGATTTACTGAANIAACANTGTAAATTCCCTGAGGT 720	95.4%	DB 2;	Length 1032;
QY	661	TTTTAACTCACCATTGATTTACTGAANIAACANTGTAAATTCCCTGAGGT 720	QY	661 TTTTAACTCACCATTGATTTACTGAANIAACANTGTAAATTCCCTGAGGT 720	96.8%	Pred. No. 1.6e-238;	
Db	661	TTTTAACTCACCATTGATTTACTGAANIAACANTGTAAATTCCCTGAGGT 720	Db	661 TTTTAACTCACCATTGATTTACTGAANIAACANTGTAAATTCCCTGAGGT 720	99%	Indels 5;	Gaps 1;
QY	721	TATCTACTGTATAAAGTAATAATCATATCAATAATCTGGGGATCATGCC 780	QY	721 TATCTACTGTATAAAGTAATAATCATATCAATAATCTGGGGATCATGCC 780	99	Conservative Matches	
Db	721	TATCTACTGTATAAAGTAATAATCATATCAATAATCTGGGGATCATGCC 780	Db	721 TATCTACTGTATAAAGTAATAATCATATCAATAATCTGGGGATCATGCC 780	99	Best Local Similarity	
QY	781	AGAGATTGTGGGGAGGAATGTTACAACGGTTCAATTGAAATTAAATC 840	QY	781 AGAGATTGTGGGGAGGAATGTTACAACGGTTCAATTGAAATTAAATC 840	99	Matches	
Db	781	AGAGATTGTGGGGAGGAATGTTACAACGGTTCAATTGAAATTAAATC 840	Db	781 AGAGATTGTGGGGAGGAATGTTACAACGGTTCAATTGAAATTAAATC 840	99	Score	
QY	841	ATTCTCTCAGAAAATCAATAACTTGGATTTTATTCCTAAACATTTAAAA 900	QY	841 ATTCTCTCAGAAAATCAATAACTTGGATTTTATTCCTAAACATTTAAAA 900	99	958;	DB 2;
Db	841	ATTCTCTCAGAAAATCAATAACTTGGATTTTATTCCTAAACATTTAAAA 900	Db	841 ATTCTCTCAGAAAATCAATAACTTGGATTTTATTCCTAAACATTTAAAA 900	99	Pred. No.	1.6e-238;
QY	901	CCACTGTAGATGTGATAAGGFACTGTGCACTTGTGACATATACATAAATTA 960	QY	901 CCACTGTAGATGTGATAAGGFACTGTGCACTTGTGACATATACATAAATTA 960	99	999;	Indels 5;
Db	901	CCACTGTAGATGTGATAAGGFACTGTGCACTTGTGACATATACATAAATTA 960	Db	901 CCACTGTAGATGTGATAAGGFACTGTGCACTTGTGACATATACATAAATTA 960	99	0;	Mismatches
QY	961	TTAAAAGTCAATCGATTCACATCTTTACACTAAAGGC 1044	QY	961 TTAAAAGTCAATCGATTCACATCTTTACACTAAAGGC 1044	99	0;	DB 2;
Db	961	TTAAAAGTCAATCGATTCACATCTTTACACTAAAGGC 1044	Db	961 TTAAAAGTCAATCGATTCACATCTTTACACTAAAGGC 1044	99	0;	Length 1032;
<hr/>							
RESULT 5							
AAQ81005			QY	361 GCACTCAACAGGAAATGGAGATGGAGATGGTTCACTGTGTAAGCAGGATGGCTC 420			
ID AAQ81005		standard; DNA; 1032 BP.	Db	361 GCAGCTCAACAGGAAATGGAGATGGTTCACTGTGTAAGCAGGATGGCTC 420			
XX			QY	361 GCACTCAACAGGAAATGGAGATGGTTCACTGTGTAAGCAGGATGGCTC 420			
AC			Db	361 GCACTCAACAGGAAATGGAGATGGTTCACTGTGTAAGCAGGATGGCTC 420			
XX			QY	421 GATCCTCTGACCTCGTGTATCGGCCCTGATGAGGGAGTGGATTCAGCGAT 480			
DT 25-MAR-2003	(revised)		Db	421 GATCCTCTGACCTCGTGTATCGGCCCTGATGAGGGAGTGGATTCAGCGAT 480			
DT 22-AUG-1995	(first entry)		QY	421 AAGAGGGCACAGCTGTGCTTCATCTCTGAGGTGTGCACTGAGGTGGAGAGC 360			
DE BAGE tumor rejection antigen precursor DNA.			Db	421 AAGAGGGCACAGCTGTGCTTCATCTCTGAGGTGTGCACTGAGGTGGAGAGC 360			
XX			QY	361 GCACTCAACAGGAAATGGAGATGGAGATGGAGATGGCTCAGTCTCC 600			
KW BAGE; tumor rejection antigen precursor; diagnosis; HLA; human leukocyte antigen MHC; major histocompatibility complex; TRAP; cancer; melanoma; ss.			Db	361 TCTATCATTCATCCCTGTGCAAGGCCCTCCCTTCAGGATTCAGCTC 600			
OS Homo sapiens.			QY	541 TGCTTGTCAGAACACATGACCAAGCTCTGAAGAGTGAACCTTACAGTAGAC 660			
XX			Db	541 TCTATCATTCATCCCTGTGCAAGGCCCTCCCTTCAGGATTCAGCTC 600			
PN W09500159-A1.			QY	601 TGCTTGTCAGAACACATGACCAAGCTCTGAAGAGTGAACCTTACAGTAGAC 660			
XX			Db	601 TCCTTGTCAGAACACATGACCAAGCTCTGAAGAGTGAACCTTACAGTAGAC 660			
PD 05-JAN-1995.			QY	661 TTTTAAACTTCACCAATGATTACTGAAATTAACAAATGTTGAAATTCCCTGAGTGT 720			
XX			Db	661 TTTTAAACTTCACCAATGATTACTGAAATTAACAAATGTTGAAATTCCCTGAGTGT 720			
PF 10-JUN-1994;	94NO-US006534.		QY	721 TATTCTACTGTATTAAAGGTAATAATCATATCAAATGTTGAAATTCCCTGAGTGT 720			
XX			Db	721 TATTCTACTGTATTAAAGGTAATAATCATATCAAATGTTGAAATTCCCTGAGTGT 720			
PR 17-JUN-1993;	93US-00079110.		QY	781 AGAGATTGTGGGGGGGG-----			
PR 15-FEB-1994;	94US-00196630.		Db	781 AGAGATTGTGGGGGGGG-----			
XX			QY	873 GTTTTATTCTTAAACATTTAAACACTGTAGATGTTAAATGGACTCTGTC 932			
(LUDWIG) LUDWIG INST CANCER RES.			Db	901 GTTTTATTCTTAAACATTTAAACACTGTAGATGTTAAATGGACTCTGTC 960			
Van Der Bruggen P, Boon-Falleur T, Coulie P, Renaud J;			QY	933 AGTATTCTGACATATACATTAATTTAAAGTCATCACTATTCAACATCTTTA 992			
XX			Db	961 AGTATTCTGACATATACATTAATTTAAAGTCATCACTATTCAACATCTTTA 1020			
PS Claim 1; Page 18; 33pp; English.			QY				
XX			PT The DNA encoding the tumor rejection antigen precursor BAGE was isolated from a cDNA library prepared from RNA from human melanoma cell line M22-MEL. The sequence and its encoded protein and derived peptide may be used in the diagnosis and therapy of cancers, e.g. melanomas. (Updated on 25-MAR-2003 to correct PN field.)				
PT Sequence 1032 BP; 283 A; 212 C; 254 G; 283 T; 0 U; 0 Other;			XX				

QY 993 CACTAAAGGCC 1004  
 QY 1021 CACTAAAGGCC 1032

**RESULT 6**  
 ABU55839/c  
 ID ABU55839 standard; DNA; 11162 BP.  
 XX ABL55839;  
 AC  
 XX DT 15-JUL-2002 (first entry)  
 DE Human mutant tissue-type plasminogen activator.  
 XX PR 04-SEP-2000; 2000WO-CN000260.  
 XX PA (HUNA-) HUNAN ROYAL BIOTECH.  
 XX PI Xia J;  
 XX DR; 2002-292268/33.  
 XX PS Disclosure; Page 18-23; 27pp; Chinese.  
 XX CC The sequence represents the mutant human tissue-type plasminogen activator used in the invention. The invention relates to a novel cell line expressing mutant human tissue-type plasminogen activator comprising an Accession Number of CCRCC C200006. The cell line of the invention has thrombolytic activity. The method is for the construction of a cell line expressing mutant human tissue-type plasminogen activator, e.g. in use for treating thrombosis by local dissolution of thrombus in blood vessels.  
 XX SQ Sequence 11162 BP; 2988 A; 2679 C; 2555 G; 2940 T; 0 U; 0 Other;

Query Match 52.2%; Score 524.4; DB 6; Length 11162;  
 Best Local Similarity 98.9%; Pred. No 2.1e-125.  
 Matches 528; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

QY 471 TTACAGCGATGTGCAATTGAACTTGGACCATATGGCAATTCATAATGGCTGAAAGAGAAA 530  
 Db 6578 TTTCAGCAATGCGATTTTAAGCACTTGGCCACTATCAAATGCTGTGAGAGAA 6519  
 QY 531 TGTACCCAGATGTATCATATCCTGTGGTGAGAGGGGCTTCAAGATTTCAGT 590  
 Db 6518 TGTACCCAGATGTATCATATCCTGTGGTGAGAGGGGCTTCAAGATTTCAGT 6459  
 QY 591 CACATCTTCTGCTTGTGAGAACATGACCAATGCAAGGTGAATTT 710  
 Db 6458 CACATCTTCTGCTTGTGAGAACATGACCAAGTCTGAGTGAATTT 6399  
 QY 651 AGCGATGAGACTTTAAACTCAACCAATGATTACTGAAATACAAATGGTGAATTT 710  
 Db 6398 AGGGATGAGCTTAAACTCAACCAATGATTACTGAAATACAAATGGTGAATTT 6339  
 QY 711 CCCTGATGTTTATCTACTGTATTAAGGTAATAATCATAAATCTGAGG 770  
 Db 6338 CCCTGATGTTTATCTACTGTATTAAGGTAATAATCATAAATCTGAGG 6279

**RESULT 7**  
 ABL5504/c  
 ID ABL5504 standard; DNA; 11162 BP.  
 XX AC ABL5504;  
 XX DT 27-MAY-2002 (first entry)  
 DE Gene vector sequence SEQ ID NO 2.  
 XX PR Human; D chromosome; G chromosome; gene therapy; nucleolus; TPA; FIX;  
 XX PA (XIAA-) XIA J.  
 XX PI Unidentified.  
 XX DR; 2002-291858/33.  
 XX PS Disclosure; Page 69-74; 76pp; Chinese.

XX CC The invention relates to a human source gene leading sequence of target gene from DNA sequences in the short arms of the human D or G group chromosome, which does not have any important physiological function. The leading sequence is useful in gene therapy for targeting a gene in a vector for transfer into nucleolus of cells for active expression. The sequence may be used for the production of desirable proteins for treating, e.g. thrombotic diseases and haemophilia B with TPA and FIX, respectively. DNA sequences such as those in the short arms of the human D or G group chromosome do not have any important physiological function and are safely used in gene therapy. The present sequence is that of a gene vector sequence, useful to the invention.

XX SQ Sequence 11162 BP; 2988 A; 2679 C; 2555 G; 2940 T; 0 U; 0 Other;

Query Match 52.2%; Score 524.4; DB 6; Length 11162;  
 Best Local Similarity 98.9%; Pred. No 2.1e-125;  
 Matches 528; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 471 TTACAGCGATGTGCAATTGAACTTGGACCATATGGCAATTCATAATGGCTGAAAGAGAAA 530

Db	6578	TTCAGCAAATGCAATTGTAAAGCACCTTGGACCACTATAAATGGCTGTGAAGAGAAA	6519	CC DNA sequences having not less than 50% homology with the above sequence, and does not have any important physiological functions. The invention also relates to a gene vector and gene expression strategy. The leading sequence is useful in gene therapy and production of protein e.g. for treating tumour and genetic diseases. This polynucleotide sequence represents the DNA of the gene vector sequence (pGEM) of the invention
QY	531	TGTACCCAGATGATCATATCCTTGTGTGATGCCGCCTTCAAGATTTCACT	590	CC XX
Db	6518	TGTACCCAGATGATCATATCCTTGTGTGATGCCGCCTTCAAGATTTCACT	6459	CC XX
QY	591	CACATCTCCGTCTTGTCAGAACACATGACCAAGCTCTGAAGATGTAAGTTACT	650	CC XX
Db	6458	CACATCTCCGTCTTGTCAGAACACATGACCAAGCTCTGAAGATGTAAGTTACT	6399	XX
QY	651	ACGATAGACTTTAACCTCAACATGATTACTGAAATAATCATTAATCTGAGG	770	Query
Db	6398	ACGATAGACTTTAACCTCAACATGATTACTGAAATAATCATTAATCTGAGG	6279	Match 52%; Score 524.4; Length 11162;
Db	6338	CCCTGAGTGTATTCTACTGTATAAAGGTATAATCATTAATCTGAGG	6279	Best Local Similarity 98.9%; Pred. No. 2.1e-125;
QY	771	GATCATTGCCAGAGATTGTTGGGAGGAATGTTTACCGGTTTCAATTGAAATTAA	830	Matches 6; Indels 0; Gaps 0;
Db	6278	GATCATTGCCAGAGATTGTTGGGAGGAATGTTTACCGGTTTCAATTGAAATTAA	6219	QY
QY	831	CCAAAAGTTATTCTCGAAAAATCAATAAATGGTTGCAAGTTTATCTTAAAC	890	Db
Db	6218	CCAAAAGTTATTCTCGAAAAATCAATAAATGGTTGCAAGTTTATCTTAAAC	6159	QY
QY	891	ATTTAAAAACCACTGTAAGATGATGTAATAAGGACTCTGCAATTTCTGACATAC	950	Db
Db	6158	ATTTAAAAACCACTGTAAGATGATGTAATAAGGACTCTGCAATTTCTGACATAC	6099	QY
QY	951	TATAAAATTATAAAAGTCAATGATTTCAACATCTTACACTAAAGGCC	1004	Db
Db	6098	TATAAAATTATAAAAGTCAATGATTTCAACATCTTACACTAAAGGCC	6045	QY
RESULT 8				Db
AK9658/c				6278 GATGTTGCCAGAGATTGGGGAAATGTTACGGTTTCATGAAATTAAAT
ID	A99658	standard; DNA; 11162 BP.		6219
XX				QY
AAK9658;				831 CCAAAAGTATTCTCAGAAAATCAATAAAGTTGCAATTTCATGAAATTAAAT
AC				6218 CCAAAAGTATTCTCAGAAAATCAATAAAGTTGCAATTTCATGAAATTAAAT
XX				6159
DT	08-JUL-2002	(first entry)		QY
XX				891 ATTAAAACCCTGAGATGATGTAATAAGGACTCTGCAATTTCATGAAATTAAAT
DE		DNA of the gene vector sequence (pGEM).		950
XX				Db
KW		Cyostatic; human; gene leading sequence; short arm D group; tumour; gene therapy; short arm G group; genetic disease; gene vector sequence; pGEM; ds.		6158 ATTAAAACCCTGAGATGATGTAATAAGGACTCTGCAATTTCATGAAATTAAAT
KW				6099
XX				QY
OS		Homo sapiens.		951 TATAAAATTAAAGTCAATGATGATGTAATAAGTCAACATCTTACACTAAAGGCC
PN		WO200220803-A1.		1004
XX				Db
PD	14-MAR-2002.			6045
XX				QY
PP	16-FEB-2001;	2001WO-CN000126.		950
XX				Db
PR	17-JUL-2000;	2000WO-CN000203.		6099
XX				XX
PA	(XIAJ/ ) XIA J.			XX
XX				XX
PI	Xia J;			XX
XX				XX
DR	WPI; 2002-304383/34.			XX
XX				XX
PT	Human source chromosome-originated leading sequence as target gene in vector for expression, useful in gene therapy and production of protein e.g. for treating tumor and genetic diseases.			XX
PT				XX
XX				XX
PS	Example 2; Page 77-82; 86pp; Chinese.			XX
XX				XX
CC	The invention relates to a human source gene leading sequence selected from a DNA sequence in the chromosomal short arm of D and G groups, or			CC









This Page Blank (uspto)

Copyright GenCore version 5.1.6  
(c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 9, 2004, 05:43:15 ; Search time 105 Seconds

(without alignments)  
5306.388 Million cell updates/sec

Title: US-10-081-108-1

Perfect score: 1004

Sequence: 1 CGCCAAATTAGGGTCTCGG.....ATCTTTACACTAAAAAGGC 1004

Scoring table: IDENTITY NUC Gapext 1.0

Searched: Gapop 10.0 , Gapext 1.0

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:\*

1: /cgcn2\_6/.ptodata/2/ina/SA\_COMB.seq:\*

2: /cgcn2\_6/.ptodata/2/ina/5B\_COMB.seq:\*

3: /cgcn2\_6/.ptodata/2/ina/6A\_COMB.seq:\*

4: /cgcn2\_6/.ptodata/2/ina/6B\_COMB.seq:\*

5: /cgcn2\_6/.ptodata/2/ina/BCTUS\_COMB.seq:\*

6: /cgcn2\_6/.ptodata/2/ina/backtus1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	1004	100.0	1004	2	US-08-389-360-1	Sequence 1, Appli	
2	1004	100.0	1004	3	US-09-038-328-1	Sequence 1, Appli	
3	1004	100.0	1004	3	US-09-056-105-10	Sequence 20, Appli	
4	1004	100.0	1004	4	US-09-435-524-1	Sequence 1, Appli	
5	1004	100.0	1004	4	US-09-1382-497-1	Sequence 1, Appli	
6	96.6	96.2	1032	1	US-08-079-1110A-1	Sequence 1, Appli	
7	96.6	96.2	1032	1	US-08-196-630A-1	Sequence 1, Appli	
8	96.6	96.2	1032	3	US-08-573-186-1	Sequence 1, Appli	
9	83	8.3	39200	4	US-10-027-983-11	Sequence 11, Appli	
10	82.4	8.2	859	4	US-09-535-008-18	Sequence 58, Appli	
11	82.4	8.2	4698	4	US-09-439-261-34	Sequence 24, Appli	
12	82.4	8.2	4698	4	US-09-227-613-33	Sequence 33, Appli	
c 13	82.4	8.2	15602	4	US-09-844-634-17	Sequence 17, Appli	
c 14	82.4	8.2	35100	1	US-08-306-691B-19	Sequence 19, Appli	
c 15	82.4	8.2	35100	5	PCT-US93-06251-19	Sequence 19, Appli	
c 16	82.2	8.2	39982	4	US-09-820-924-3	Sequence 3, Appli	
c 17	82.4	8.2	84495	4	US-09-797-906-3	Sequence 3, Appli	
c 18	82.4	8.2	319608	4	US-09-539-333D-1	Sequence 1, Appli	
c 19	82.4	8.2	319608	4	US-09-679-409-1	Sequence 1, Appli	
c 20	82.2	8.2	7210	2	US-08-257-963B-10	Sequence 10, Appli	
c 21	82.2	8.2	7210	4	US-08-367-841A-10	Sequence 10, Appli	
c 22	82.2	8.2	7210	5	PCT-US95-07201-10	Sequence 10, Appli	
c 23	82.2	8.2	14581	4	US-08-520-373D-4	Sequence 4, Appli	
c 24	82.2	8.2	22481	4	US-08-367-841A-43	Sequence 43, Appli	
c 25	82.2	8.2	22481	5	PCT-US95-07201-43	Sequence 2, Appli	
c 26	82.2	8.2	22484	4	US-09-875-223-2	Sequence 2, Appli	
c 27	82.2	8.2	22484	4	US-09-875-114-2	Sequence 3, Appli	

%

^

RESULT 1  
US-08-389-360-1  
Sequence 1, Application US/08389360  
Patent No. 587017  
GENERAL INFORMATION:  
APPLICANT: van der Bruggen et al.  
TITLE OF INVENTION: ISOLATED PEPTIDE WHICH FORMS COMPLEXES WITH MHC MOLECULE HLA-Cw\*1601 AND USES NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Felfe & Lynch  
STREET: 805 Third Avenue  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/389,360  
FILING DATE: Herewith  
CLASSIFICATION: 436  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/196,630  
FILING DATE: February 15, 1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/079,110  
FILING DATE: June 17, 1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Pasqualini, Patricia A.  
REGISTRATION NUMBER: 34,894  
REDOCKET NUMBER: LUD 5310.2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-3884  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1004  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-389-360-1

Query Match Score 100.0%; Best Local Similarity 100.0%; Matches 1004; Conservative 0; Missmatches 0; Indexes 0; Gaps 0;

Patent No. 6110694  
 GENERAL INFORMATION:  
 APPLICANT: van der Bruggen et al.  
 TITLE OF INVENTION: ISOLATED PEPTIDE WHICH FORMS COMPLEXES WITH MHC MOLECULE HLA-Cw\*1601 AND USES NUMBER OF SEQUENCE: 9  
 TITLE OF INVENTION: THEREOF  
 CORRESPONDENCE ADDRESS:  
 STREET: 805 Third Avenue  
 CITY: New York  
 STATE: New York  
 COUNTRY: USA  
 ZIP: 10022  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette  
 COMPUTER: IBM PS/2  
 OPERATING SYSTEM: PC-DOS  
 SOFTWARE: Wordperfect  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/038,328  
 FILING DATE:  
 CLASSIFICATION:  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: 08/389,360  
 FILING DATE:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/079,110  
 FILING DATE: June 17, 1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Pasqualini, Patricia A.  
 REGISTRATION NUMBER: 34,994  
 REFERENCE/DOCKET NUMBER: LUD 5310.2  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 688-9200  
 TELEFAX: (212) 638-3884  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1004;  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 US-09-038-328-1

Query 1 CGCCATTAGGGCTCCGATATCCCGTCTGGCTGAGCTAGGACC 60  
 Db 1 CGCCATTAGGGTCTCCGTAATCCCGTCTGGCTGAGGACC 60  
 Query 61 AGGAACCGGGACTGGGTGAGCTGAACTGGCTGAGCTGATG 120  
 Db 61 AGGAAGGGGAACTGGGTGAGCTGAACTGGCTGAGCTGATG 120  
 Query 121 GTGGTGCACAGAGATGGCAGGGCTGGAGCTGTTAGGAGTGG 180  
 Db 121 GTGGTGCACAGAGATGGCAGGGCTGGAGCTGTTAGGAGTGG 180  
 Query 121 GTGGTGCACAGAGATGGCAGGGCTGGAGCTGTTAGGAGTGG 180  
 Db 121 GTGGTGCACAGAGATGGCAGGGCTGGAGCTGTTAGGAGTGG 180  
 Query 181 AGAGATTGGGGGGAAATGGTATCAACGGTTCATGAAATTCAAAAGTT 840  
 Db 781 TATTCTACTGTTAATAAAGGTAAATACATAATTCATGAAATTTCCTGAGTT 720  
 Query 721 TATTCTACTGTTAATAAAGGTAAATACATAATTCATGAAATTTCCTGAGTC 780  
 Db 661 TTTAAACTCAACCAATGTTTAAATCAAAATGTTTAAATCAAAATGTT 720  
 Query 781 AGAGATTGGGGGGGGAAATGGTATCAACGGTTCATGAAATTCAAAAGTT 840  
 Db 781 AGAGATTGGGGGGAAATGGTATCAACGGTTCATGAAATTCAAAAGTT 840  
 Query 841 ATTCCTCAGAAAATCAAAATCAAAATGTTGCAATTCTTAACATTTAAAA 900  
 Db 841 ATTCCTCAGAAAATCAAAATCAAAATGTTGCAATTCTTAACATTTAAAA 900  
 Query 901 CCACCTGATGATGATGAAATAGGAACTGGCAGTATTCTGAGTATTCTG 960  
 Db 901 CCACCTGATGATGATGAAATAGGAACTGGCAGTATTCTGAGTATTCTG 960  
 Query 961 TTAAAAAGTCATGATCAACATCTTTACACTAAAGGCC 1004  
 Db 961 TTAAAAAGTCATGATCAACATCTTTACACTAAAGGCC 1004  
 Query 301 AGAACCCACAGCTCTGCTTCACTCTGAGGTTGTCAGCCAGGAGC 360  
 Db 301 AGAACCCACAGCTCTGCTTCACTCTGAGGTTGTCAGCCAGGAGC 360  
 Query 361 GCAGCTAACAGGAGAAATAGGAGATGAGGAGCTGTCAGGATGAGCTG 300  
 Db 361 GCAGCTAACAGGAGAAATAGGAGATGAGGAGCTGTCAGGATGAGCTG 300

RESULT 2  
 US-09-038-328-1, Application US/09038328  
 Sequence 1,

Db 361 GCAGCTAACRGGCCATTAGGAGAGATTCACTGGTCAAGCAGATGGTC 420  
 Qy 421 GATCCCTGAACTCGTATGCCCTTCAAAGTGCCGAGATTACGGCAT 480  
 Db 421 GATCCCTGAACTCGTATGCCCTTCAAAGTGCCGAGATTACGGCAT 480  
 Qy 481 GTGCATTTCGAAGAACATTGGAGOCACATCAATGCTGTGAAAGAAAATGACCC2GA 540  
 Db 481 GTGCATTTCGAAGAACATTGGAGOCACATCAATGCTGTGAAAGAAAATGACCC2GA 540  
 Qy 541 TGTATCATTATCCTTGTGCTCAGGCCGCTCTTCTAGGATTCTGTCATCTCC 600  
 Db 541 TGTATCATTATCCTTGTGCTCAGGCCGCTCTTCTAGGATTCTGTCATCTCC 600  
 Qy 601 TGCTTGTCAGAAACATGACCAAGCTGAAAGAATGTAAGTTACCCATAC 660  
 Db 601 TGCTTGTCAGAAACATGACCAAGCTGAAAGAATGTAAGTTACCCATAC 660  
 Qy 661 TTAAACCTCAACAAATGATTAATCTGAAATAACAAATGATTAATCTGAGNGT 720  
 Db 661 TTAAACCTCAACAAATGATTAATCTGAAATAACAAATGATTAATCTGAGNGT 720  
 Qy 721 TATTCTACTGTATAAACGTTAAATACATTAATCATTAATCATTAATCTGGGGATCATTCGCC 780  
 Db 721 TATTCTACTGTATAAACGTTAAATACATTAATCATTAATCATTAATCTGGGGATCATTCGCC 780  
 Qy 781 AGAGATTGTTGGGCCAAATGTCAACGGTTTCAATTGAAATTTACCAAAAGTT 840  
 Db 781 AGAGATTGTTGGGCCAAATGTCAACGGTTTCAATTGAAATTTACCAAAAGTT 840  
 Qy 841 ATTCTCTGAAATAATCAAAATAAAGTTGCACTGTTTATCTTAACATTTAA 900  
 Db 841 ATTCTCTGAAATAATCAAAATAAAGTTGCACTGTTTATCTTAACATTTAA 900  
 Qy 901 CCACTGTAAATGATGTAAATAGGACTGTGCAATTTCTGCATATAAAATA 960  
 Db 901 CCACTGTAAATGATGTAAATAGGACTGTGCAATTTCTGCATATAAAATA 960  
 Qy 961 TAAAGTCAATCAGTATCAACATCTTACACTAAAGCC 1004  
 Db 961 TAAAGTCAATCAGTATCAACATCTTACACTAAAGCC 1004

RESULT 3  
 US-10-05-105-20  
 Sequence No. 20, Application US/09056105  
 ; Patent No. 6287569  
 ; GENERAL INFORMATION:  
 ; APPLICANT: KIPPS, THOMAS J.  
 ; APPLICANT: WU, YUNQI  
 ; TITLE OF INVENTION: VACCINES WITH ENHANCED INTRACELLULAR  
 ; FILE REFERENCE: 233/221  
 ; CURRENT APPLICATION NUMBER: US/09/056,105  
 ; EARLIER APPLICATION NUMBER: 60/0043,467  
 ; EARLIER FILING DATE: 1997-04-10  
 ; NUMBER OF SEQ ID NOS: 35  
 ; SOFTWARE: FastSEQ For Windows Version 3.0  
 ; SEQ ID NO: 20  
 ; LENGTH: 1004  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-09-056-105-20

Query Match 100.0%; Score 1004; DB 3; Length 1004;  
 Best Local Similarity 100.0%; Pred. No. 3.7e-58;  
 Matches 1004; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCCAAATTAGGCTCTCCCGTATCTCCCGTAGCTGAGCTGAGCC 60  
 Db 1 CGCCAAATTAGGCTCTCCCGTATCTCCCGTAGCTGAGCTGAGCC 60

RESULT 4  
 US-09-435-524-1  
 ; Sequence 1, Application US/09435524  
 ; Patent No. 6465134  
 ; GENERAL INFORMATION:  
 ; APPLICANT: van der Bruggen et al.  
 ; TITLE OF INVENTION: ISOLATED PEPTIDE WHICH FORMS COMPLEXES

WITH MHC MOLECULE HLA-Cw\*1601 AND USES

THEREOF

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pelfe & Lynch  
STREET: 805 Third Avenue  
CITY: New York

COUNTRY: USA

ZIP: 10022

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette  
COMPUTER: IBM PS/2

OPERATING SYSTEM: PC-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/435,524

FILING DATE: 08 -No. 645184-1999

CLASSIFICATION: &lt;Unknown&gt;

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/038,328

FILING DATE: &lt;Unknown&gt;

APPLICATION NUMBER: US/079,110

FILING DATE: June 17, 1993

ATTORNEY/AGENT INFORMATION:

NAME: Paganini, Patricia A.

REGISTRATION NUMBER: 34,894

REFERENCE DOCKET NUMBER: LUD 5310.2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 658-9200

TELEFAX: (212) 838-3884

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US -09-435-524-1

Query Match 100.0% Score 1004; DB 4; Length 1004;  
Best Local Similarity 100.0% Pred. No. 3\_7e-28; Indels 0; Gaps 0;

Matches 1004; Conservative 0; Mismatches 0;

Qy 1 CGCCATTAGGGCTCCCTATCCCTGAGCTGCCTGAGGCC 60

Db 1 CGCCATTAGGGCTCCCTGAGCTGCCTGAGGCC 60

Qy 1 AGGAAAGGGAGCTGGGGCTGAGCTGTACACCGTGTCACTCTGATG 120

Db 1 AGGAGAGGGAGCTGGGGCTGAGCTGTACACCGTGTCACTCTGATG 120

Qy 121 GTGGTGGCACAGACATGGCAGCAGTGGAGTGGAGCTGAGGGCTGAGCTG 180

Db 121 GTGGTGGCACAGACATGGCAGCAGTGGAGCTGAGGGCTGAGCTG 180

Qy 181 AGTGGGGCTGAGCTGAGATGGGGCAAGGGTTTCGGCATTTGTGCCAGC 240

Db 181 AGTGGGGCTGAGCTGAGATGGGGCAAGGGTTTCGGCATTTGTGCCAGC 240

Qy 241 TGCTCCAAGCAGGGTGAAGGGGACTCCCTGTTGAGCTGGAGCTG 300

Db 241 TGCTCCAAGCAGGGTGAAGGGGACTCCCTGTTGAGCTGGAGCTG 300

Qy 301 AAGACGGCACGCTGCTGCTGAGCTGGCTGAGCTGGAGCTG 360

Db 301 AAGACGGCACGCTGCTGAGCTGGCTGAGCTGGAGCTG 360

Qy 361 GCAGCTAACAGGCAATTAGGAGAATAGGAGAATGAGATGGCTC 420

Db 361 GCAGCTAACAGGCAATTAGGAGAATGAGATGGCTC 420

Qy 421 GATCTCTGACCTCGTGAATCGCCGGCTTGGCCATTACAGGAT 480

## RESULT 5

US-09-382-497-1

; Sequence 1, Application US/09382497

;

; Parent No. 6638512

; GENERAL INFORMATION:

; APPLICANT: van der Bruggen et al.

; TITLE OF INVENTION: ISOLATED PEPTIDE WHICH FORMS COMPLEXES

; WITH MHC MOLECULE HLA-Cw\*1601 AND USES

; TITLE OF INVENTION: THEREOF

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fife &amp; Lynch

; STREET: 805 Third Avenue

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10022

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette

; COMPUTER: IBM PS/2

; OPERATING SYSTEM: PC-DOS

; SOFTWARE: Wordperfect

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/382,497

; FILING DATE:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/389,360

; FILING DATE:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/079,110

; FILING DATE: June 17, 1993

; ATTORNEY/AGENT INFORMATION:

NAME: Pasqualini, Patricia A.  
 REFERENCE/DOCKET NUMBER: 34,994  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 688-9200  
 TELEFAX: (212) 838-3884  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1004  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 US-09-382-497-1

Query Match 100.0%; Score 1004; DB 4; Length 1004;  
 Best Local Similarity 100%; Pred. No. 3,7-238; Pairs 0;  
 Matches 1004; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCCZATTAGGGCTCCCGTATCCTCCCTGACTGTGCTGAGGACC 60  
 Db 1 CGCCAATTAGGGCTCCCGTATCCTCCCTGACTGTGCTGAGGACC 60

QY 61 AGGAAAGGGGACTGGGGCTGAGCTGTAAACCCGACTCGTGTCACTCTGGATG 120  
 Db 61 AGGAAAGGGGAGTGAGGTGGCTGGAGCTGTAAACCTGGCTGTTCACTCTGGATG 120

QY 121 GTGGTGGCAAAGAGATGCCAGCTGGAGCTGTAGGAGGGCCCTGAGCGTAGG 180  
 Db 121 GTGGTGGCACAGAGATGCCAGCTGGAGGTAGGAGGGCTGAGCGTAGG 180

QY 181 AGTGGGGCTGGAGGATAAGATGCCAGCTGGAGCTGTAGGAGGGCTGAGCGTAGG 240  
 Db 181 AGTGGGGCTGGAGGATAAGATGCCAGCTGGAGCTGTAGGAGGGCTGAGCGTAGG 240

QY 241 TGCTCCAAGCCAGCTGTGAAGAGGGTCCCTGTGGTGACTGTGAGGGTGGAGCTG 300  
 Db 241 TGCTCCAAGCCAGCTGTGAAGAGGGTCCCTGTGGTGACTGTGAGGGTGGAGCTG 300

QY 301 AAGGGGCAAGCTGTGCTCATCTGTGAGGTGTGGAGCCACTGTGATGGAGC 360  
 Db 301 AAGGGGCAAGCTGTGCTCATCTGTGAGGTGTGGAGCCACTGTGATGGAGC 360

QY 361 GCAGTCACAGGGCAATTAGGGGATGAGGTGACTGTGCTGAGGATGGCTC 420  
 Db 361 GCAGTCACAGGGCAATTAGGGGATGAGGTGACTGTGCTGAGGATGGCTC 420

QY 421 GATCCCTTAACCCGTGATGCCGCTTGCCTCATCTGTGGCTTCCAAAGTCCGAT 480  
 Db 421 GATCCCTAACCCGTGATGCCGCTTGCCTCATCTGTGGCTTCCAAAGTCCGAT 480

QY 481 GTGGATTGTGAACTTGGGCCACTATCAATGTGTAAAGAATGTACCGA 540  
 Db 481 GTGGATTGTGAACTTGGGCCACTATCAATGTGTAAAGAATGTACCGA 540

QY 541 TGPATCATATCCTTGTGAGCACTTGGGCCACTATCAATGTGTAAAGAATGTACCGA 600  
 Db 541 TGPATCATATCCTTGTGAGCACTTGGGCCACTATCAATGTGTAAAGAATGTACCGA 600

QY 601 TGTCTTGCAAGACATGACAAAGCTCTGAGCTGAGCTGATGTTAATAGCTAGCATAGC 660  
 Db 601 TGTCTTGCAAGACATGACAAAGCTCTGAGCTGAGCTGATGTTAATAGCTAGCATAGC 660

QY 661 TTAAATCTTCAACCAATGTTTACGAAATAACAAATGTGTAAATTCCCTGAGTGT 720  
 Db 661 TTAAATCTTCAACCAATGTTTACGAAATAACAAATGTGTAAATTCCCTGAGTGT 720

QY 721 TAATCTCATCTGTAAAGGTAAATATCAATCATTAATCTGAGGATCATGGCC 780  
 Db 721 TAATCTCATCTGTAAAGGTAAATCAATCATTAATCTGAGGATCATGGCC 780

QY 781 AGAGATTTGGAGGGGAAAGTTTACCGTTTATGAAATAATCCAAAGT 840  
 Db 781 AGAGATTTGGAGGGGAAAGTTTACCGTTTATGAAATAATCCAAAGT 840

QY 841 ATTCTCTCAGAAAATCAAATAAGTTGGATGTTTATTCTAAACATTAAAAA 900  
 Db 841 ATTCTCTCAGAAAATCAAATAAGTTGGATGTTTATTCTAAACATTAAAAA 900

QY 901 CCACTGTAGATGATGTTAAATAGGGACTGTGCAATTCTGACATACTATAAAATTA 960  
 Db 901 CCACTGTAGATGATGTTAAATAGGGACTGTGCAATTCTGACATACTATAAAATTA 960

QY 961 TTAAGAAGTCATGATGTTAACTCAACTCTTACACTAAAGGCC 1004  
 Db 961 TTAAGAAGTCATGATGTTAACTCAACTCTTACACTAAAGGCC 1004

RESULT 6  
 US-09-110A-1  
 Sequence 1, Application US/08079110A  
 Patent No. 5571111  
 GENERAL INFORMATION:  
 APPLICANT: van der Bruggen, Pierre; Boon-Palleur,  
 APPLICANT: Thierry, Coulie, Pierre; Renaud, Jean-Christophe  
 TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES  
 NUMBER OF SEQUENCES: 6  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Felfe & Lynch  
 STREET: 805 Third Avenue  
 CITY: New York City  
 STATE: New York  
 COUNTRY: USA  
 ZIP: 10022  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage  
 COMPUTER: IBM PS/2  
 OPERATING SYSTEM: PC-DOS  
 SOFTWARE: Wordperfect  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/079,110A  
 FILING DATE: 17-JUN-1993  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Hanson, No. 55171iman D.  
 REGISTRATION NUMBER: 30,946  
 REFERENCE/DOCKET NUMBER: LUD 5310  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 688-9200  
 TELEFAX: (212) 838-3884  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1032 base pairs  
 TYPE: nucleic acid  
 STANDEENESS: single  
 TOPOLOGY: linear  
 US-08-079-110A-1  
 Query Match 96.2%; Score 966; DB 1; Length 1032;  
 Best Local Similarity 97.3%; Pred. No. 5,1-e248; Pairs 0; Mismatches 0; Indels 28; Gaps 1;

QY 1 CGCCATTAGGGCTCCGGTATCTCCCGCTGAGCTGCTGCTGAGCTGAGGCC 60  
 Db 1 CGCCATTAGGGCTCCGGTATCTCCCGCTGAGCTGCTGCTGAGGCC 60

QY 61 AGGAAAGGGGGAGGTGGAGCTGAGCTGAGGCCAGCTGGCTCACCTGGATG 120  
 Db 61 AGGAAAGGGGGAGGTGGAGCTGAGGCCAGCTGGCTCACCTGGATG 120

QY 121 GTGGTGGCAAAGAGATGGAGCTGAGGCTGAGGGCTGAGCTGAGGCC 180  
 Db 121 GTGGTGGCAAAGAGATGGAGCTGAGGCTGAGCTGAGGCC 180

QY 181 AGAGATTTGGAGGGAAAGTTTACCGTTTATGAAATAATCCAAAGT 240  
 Db 181 AGAGATTTGGAGGGAAAGTTTACCGTTTATGAAATAATCCAAAGT 240





RESULT 9  
 US-10-027-983-11  
 Sequence 11, Application US/10027983  
 Patent No. 6617162  
 GENERAL INFORMATION:  
 APPLICANT: Kenneth W. Dobie  
 ATTORNEY: Mark P. Roach  
 TITLE OF INVENTION: ANTISENSE MODULATION OF ESTROGEN RECEPTOR ALPHA EXPRESSION  
 FILE REFERENCE: RTS 0340  
 CURRENT FILING DATE: 2001-12-18  
 NUMBER OF SEQ ID NOS: 98  
 SEQ ID NO 11  
 LENGTH: 392000  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: unsure  
 LOCATION: 137740  
 OTHER INFORMATION: unknown  
 NAME/KEY: unsure  
 LOCATION: 137742  
 OTHER INFORMATION: unknown  
 NAME/KEY: misc feature  
 LOCATION: (138122) .. (138221)  
 OTHER INFORMATION: n = A, T, C or G  
 NAME/KEY: unsure  
 LOCATION: 145507  
 OTHER INFORMATION: unknown  
 NAME/KEY: unsure  
 LOCATION: 151967  
 OTHER INFORMATION: n = A, T, C or G  
 NAME/KEY: misc feature  
 LOCATION: (151967) .. (1542066)  
 OTHER INFORMATION: n = A, T, C or G  
 NAME/KEY: unsure  
 LOCATION: 154217  
 OTHER INFORMATION: unknown  
 NAME/KEY: misc feature  
 LOCATION: (164037) .. (164136)  
 OTHER INFORMATION: n = A, T, C or G  
 NAME/KEY: misc feature  
 LOCATION: (174657) .. (174756)  
 OTHER INFORMATION: n = A, T, C or G  
 NAME/KEY: misc feature  
 LOCATION: (186724) .. (186323)  
 OTHER INFORMATION: n = A, T, C or G  
 NAME/KEY: misc feature  
 LOCATION: (195742) .. (195341)  
 OTHER INFORMATION: n = A, T, C or G  
 LOCATION: 202703  
 OTHER INFORMATION: unknown  
 NAME/KEY: misc feature  
 LOCATION: (203771) .. (202870)  
 OTHER INFORMATION: n = A, T, C or G  
 NAME/KEY: misc feature  
 LOCATION: (205246) .. (215602)  
 OTHER INFORMATION: n = A, T, C or G  
 NAME/KEY: misc feature  
 LOCATION: (218126) .. (218225)  
 OTHER INFORMATION: n = A, T, C or G  
 NAME/KEY: misc\_feature

LOCATION: (220360) .. (220459)  
 OTHER INFORMATION: n = A, T, C or G  
 NAME/KEY: misc feature  
 LOCATION: (222717) .. (222816)  
 OTHER INFORMATION: n = A, T, C or G  
 NAME/KEY: misc feature  
 LOCATION: (222791) .. (224030)  
 OTHER INFORMATION: n = A, T, C or G  
 NAME/KEY: misc feature  
 LOCATION: (222748) .. (227536)  
 OTHER INFORMATION: n = A, T, C or G  
 NAME/KEY: misc feature  
 LOCATION: (230157) .. (230256)  
 OTHER INFORMATION: n = A, T, C or G  
 NAME/KEY: misc feature  
 LOCATION: (231299) .. (232398)  
 OTHER INFORMATION: n = A, T, C or G  
 NAME/KEY: misc feature  
 LOCATION: (236552) .. (2366651)  
 OTHER INFORMATION: n = A, T, C or G  
 NAME/KEY: misc feature  
 LOCATION: (236552) .. (2366651)  
 OTHER INFORMATION: n = A, T, C or G  
 NAME/KEY: exon  
 LOCATION: (111288) .. (111901)  
 OTHER INFORMATION: exon 1C  
 NAME/KEY: exon-intron junction  
 LOCATION: (151129) .. (151130)  
 OTHER INFORMATION: exon 5:inttron 5  
 NAME/KEY: exon-intron junction  
 LOCATION: (299448) .. (299249)  
 OTHER INFORMATION: exon 9:inttron 9  
 NAME/KEY: exon-intron junction  
 LOCATION: (348578) .. (348579)  
 OTHER INFORMATION: exon 10:intron 10  
 NAME/KEY: intron  
 LOCATION: (348579) .. (381838)  
 OTHER INFORMATION: intron 10  
 NAME/KEY: intron:exon junction  
 LOCATION: (386185) .. (386186)  
 OTHER INFORMATION: intron 11:exon 12  
 US-10-027-983-11

Query Match 8.3%; Score 83; DB 4; Length 392000;  
 Best Local Similarity 89.9%; Pred. No. 2.4e-11; Indels 0; Gaps 0;  
 Matches 89; Conservative 0; Mismatches 0;

Qy 386 AGATGGAGTTCACTGTGTCAGCCAGATGGTCTGACTCTGATCCGGCC 445  
 Db 252283 AGACGGAGTTCACTGTGTTAGCCAGATGGCTCTGATCCGCC 252342

Qy 446 GCCTTGCCCTCCAAAGTGCCTGAGATACAGCGATGTGC 484  
 Db 252343 GCCTGAGCCCTCCAAAGTGTGGATATACGGCATGAGC 252381

RESULT 10  
 US-09-535-008-58  
 Sequence 58, Application US/09535008  
 ; Patent No. 6465639  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wong, Alexander K.C.  
 ; APPLICANT: Wong, Alexander K.C.,  
 ; ATTORNEY: Tavelgian, Sean V.  
 ; ATTORNEY: Teng, David H.-F.  
 ; TITLE OF INVENTION: BRG1 IS A TUMOR SUPPRESSOR THAT IS MUTATED IN PROSTATE  
 ; OTHER INFORMATION: AND OTHER CANCER TYPES  
 ; FILE REFERENCE: 2318-259  
 ; CURRENT FILING DATE: 2000-03-23  
 ; EARLIER APPLICATION NUMBER: US/09/535,008  
 ; CURRENT FILING DATE: 2000-03-23  
 ; EARLIER APPLICATION NUMBER: U.S. 60/125,806  
 ; EARLIER FILING DATE: 1999-03-23  
 ; NUMBER OF SEQ ID NOS: 77  
 ; SOFTWARE: Patentin Ver. 2.0



OPERATING SYSTEM: MS-DOS  
 SOFTWARE: Wordperfect 5.1  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/306,691B  
 FILING DATE: September 15, 1994  
 CLASSIFICATION: 514  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Monaco, Daniel A.  
 REGISTRATION NUMBER: 30,480  
 REFERENCE/DOCKET NUMBER: 8321-8  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (215) 568-8383  
 TELEX: (215) 568-5549  
 TELEX: No. 5734039e  
 INFORMATION FOR SEQ ID NO: 19:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 35100 base pairs  
 TYPE: nucleic acid  
 STRANDBNESS: double  
 TOPOLOGY: linear  
 US-08-306-691B-19

```

Query Match          8.2%;  Score 82.4;  DB 5;  Length 35100;
Best Local Similarity 89.0%;  Prod. No. 1..le-11;
Matches 89;  Conservative 0;  Mismatches 11;  Indels 0;  Gaps 0;
  
```

Qy 385 GAGATGGAGTTCACTGTTCAAGCAGGATGGCTCGATCTGCAGCTCTGATGCC 444

```

Db 28167 GAGATGGGTTTACCATGTTAGCCAGATGGCTCGATCTGATCACC 28226
  
```

Qy 445 CGCCCTGGCCTTCAAAGTGCCGAGATTACAGCGATGTC 484

```

Db 28227 CACCTGGCTCCAAAGTGTGAGATTACAGGCATGAGC 28266
  
```

Search completed: July 9, 2004, 10:09:37  
Job time : 110 secs

---

Query Match 8.2%; Score 82.4; DB 1; Length 35100;
 Best Local Similarity 89.0%; Prod. No. 1..le-11;
 Matches 89; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 385 GAGATGGAGTTCACTGTTCAAGCAGGATGGCTCGATCTGCAGCTCTGATGCC 444

```

Db 28167 GAGATGGGTTTACCATGTTAGCCAGATGGCTCGATCTGATCACC 28226
  
```

Qy 445 CGCCCTGGCCTTCAAAGTGCCGAGATTACAGCGATGTC 484

```

Db 28227 CACCTGGCTCCAAAGTGTGAGATTACAGGCATGAGC 28266
  
```

RESULT 15  
PCT-US93-06251-19  
Sequence 19, Application PC/TUS9306251

GENERAL INFORMATION:  
 APPLICANT: Wicksstrom, Eric and Rife, Jason P.  
 TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing Stereospecific Alkylphosphonates and Arylphosphonates  
 NUMBER OF SEQUENCES: 93  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: SCOTT, MURPHY & PRESSER  
 STREET: 400 Garden City Plaza  
 CITY: Garden City  
 STATE: NY  
 COUNTRY: USA  
 ZIP: 11530  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US93/06251  
 FILING DATE: 19930630  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: DiGiulio, Frank S.  
 REGISTRATION NUMBER: 31,346  
 REFERENCE/DOCKET NUMBER: 8586  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 516-742-4243  
 TELEFAX: 516-742-4366  
 TELEX: 230 901 SANS UR  
 INFORMATION FOR SEQ ID NO: 19:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 35100 base pairs

ALIGNMENTS

```

RESULT 1
US-10-207-655-80
; Sequence 80, Application US/10207655
; Publication No. US2003011859A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Hayden-Ledbetter, Martha S.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069_401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: Patentin version 3.0
SEQ ID NO 80
LENGTH: 1004
END

is the number of results predicted by chance to have a
similar sequence to the query sequence. It is not necessarily
the number of results that are relevant to the search.

```

**No.** is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

תְּהִלָּה

Result No.	Score	Query			DB ID	Description	
		Match	Length	DB			
1	1004	100.0	1004	15	US-10-207-655-80	Sequence 80, AppLife	
2	1004	100.0	1004	15	US-11-0-081-108-1	Sequence 1, AppLife	
C	524.4	52.2	11162	17	US-10-362-916-1	Sequence 1, AppLife	
C	524.4	52.2	13928	13	US-10-0-029-386-510	Sequence 1510, AppLife	
C	330.4	32.9	524	15	US-11-0-029-386-8714	Sequence 8714, AppLife	
C	7	164.6	345	15	US-10-0-029-386-18339	Sequence 18339, AppLife	
C	8	164.6	538	15	US-10-0-029-386-4583	Sequence 4583, AppLife	
C	9	164.2	172	15	US-11-0-029-386-22449	Sequence 22449, AppLife	
C	10	149.8	153	15	US-10-0-029-386-15211	Sequence 15211, AppLife	
C	11	149.8	62912	13	US-0-0-0-221-714A-461	Sequence 461, AppLife	
C	12	93.6	9	31516	16	US-0-0-0-087-192-1252	Sequence 1252, AppLife
C	13	88.4	8	407	13	US-0-0-0-027-632-128026	Sequence 128026, AppLife
C	14	88.0	407	13	US-0-0-0-027-632-128026	Sequence 128026, AppLife	

181	AGTGGGCGCTGGAGCAGTAAGATGGGGCCAGACCCGGTTTTCRGCATTGTCCCCAGC	240
241	TGCTCCAAGCCAGCTGTATGAAGGGAGAATGCCCTGTGTGAGGTGGAGCTG	300
241	TGCTCCAAGCCAGCTGTATGAAGGGAGAATGCCCTGTGTGAGGTGGAGCTG	300
301	AAGACGGCAAGCTCTGTGCTCATCTTCTGAGGGTGTGCCACCGGTGAGCTG	360
301	AAGACGGCAAGCTCTGTGCTCATCTTCTGAGGGTGTGCCACCGGTGAGCTG	360
361	GCGCTCAACAGGGCAATAGGGAGGATGGAGTTCACTGTGTCA	420
361	GCGCTCAACAGGGCAATAGGGAGGATGGAGTTCACTGTGTCA	420
421	GATCTCTGACTCTGTGATCGCCCGCCTTGGCCTTCCAAAGTCCCGAGATTAA	480
421	GATCTCTGACTCTGTGATCGCCCGCCTTGGCCTTCCAAAGTCCCGAGATTAA	480
481	GTGCAATTGPAAGCACTTGGGCACTATCAAATGCTGTGAGAGAAATGTA	540
481	GTGCAATTGPAAGCACTTGGGCACTATCAAATGCTGTGAGAGAAATGTA	540
541	TGATCATTAATCTTGCTTGCTCAGGGCGCTCCTTCAGGATTCACTCTTC	600
541	TGATCATTAATCTTGCTTGCTCAGGGCGCTCCTTCAGGATTCACTCTTC	600
601	TGCTTGTCCAGAACACATGGACAAAGCTGTAAAGTTACTACCGATAGAC	660
601	TGCTTGTCCAGAACACATGGACAAAGCTGTAAAGTTACTACCGATAGAC	660
661	TTTAAACTTCAACCATTGATTTACTGAAAATACAAATGTTGATAATTCTCGATG	720
661	TTTAAACTTCAACCATTGATTTACTGAAAATACAAATGTTGATAATTCTCGATG	720
721	TATTCTACTTGTATAAAAGTTATACATACTAAATCTGGGGATATTGCC	780
721	TATTCTACTTGTATAAAAGTTATACATACTAAATCTGGGGATATTGCC	780
781	AGAGATTGTGGGGAGAAATGTTATCAAAAGTTTCAATGAAATTAACTCAA	840
781	AGAGATTGTGGGGAGAAATGTTATCAAAAGTTTCAATGAAATTAACTCAA	840
841	ATTTCCTCAGAAAATCAAATGTTATCAAAAGTTTCAATGAAATTAACTCAA	900
841	ATTTCCTCAGAAAATCAAATGTTATCAAAAGTTTCAATGAAATTAACTCAA	900
901	CCACTGTGAAATGATGTTAAATAGGGCACTGTGCAATTGTTCAATGAA	960
901	CCACTGTGAAATGATGTTAAATAGGGCACTGTGCAATTGTTCAATGAA	960
961	TTAAAGTCATCAGTATCAACATCTTTCAGTAAAGGC	1004
961	TTDABANGTCATCAGTATCAACATCTTTCAGTAAAGGC	1004

RESULT 2  
-10-081-108-1  
Publication No. US20030138854A1  
GENERAL INFORMATION:  
APPLICANT: van der Bruggen et al.  
TITLE OF INVENTION: ISOLATED PEPTIDE WHICH FORMS COMPLEXES WITH MHC MOLECULE HLA-Cw\*1601 AND USES THEREOF  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Felfe & Lynch  
STREET: 805 Third Avenue  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10022

Qy	301	AGAGGGCACAGGCTTGCTTCATCTTGAGTTGCGAGAACGTTGAGAGC 360
Db	301	AAGGGCACAGGCTTGCTTCATCTTGAGTTGCGAGAACGTTGAGAGC 360
Qy	361	GCAGCTCAAAGGGACAAATAGGGAGATGGATTCACTGTGAGGATGGTCTC 420
Db	361	GCAGCTCAAAGGGACAAATAGGGAGATGGATTCACTGTGAGGATGGTCTC 420
Qy	421	GATCCTCTGACCTCTGATCGCCGCCCTGGCTTCCAAGTCGAGATTAGCGAT 480
Db	421	GATCCTCTGACCTCTGATCGCCGCCCTGGCTTCCAAGTCGAGATTAGCGAT 480
Qy	481	GTGCATTGTAAGACTTGAGGACTCATCATGCTGAGAGAAATGACCCAGA 540
Db	481	GTGCATTGTAAGACTTGAGGACTCATCATGCTGAGAGAAATGACCCAGA 540
Qy	541	TGTATCATTATCCCTGCTGAGGATTCACTCTTCAGTCACATCTTCC 600

Db 541 TGTATCATTATCCITGCTGAGGGCGCTCCTTAGGTTCACTCTTC 600  
 Qy 601 TGCTTGTCAAGAACATGACCAAGCTGCTGAAGATTAACCCATAGC 660  
 Db 601 TGCTTGTCAAGAACATGACCAAGCTGCTGAAGATTAACCCATAGC 660  
 Qy 661 TTAAACTCACCAGTATTACTGAAATAACAAATGGTAAATTACCTGC 720  
 Db 661 TTAAACTCACCAGTATTACTGAAATAACAAATGGTAAATTACCTGC 720  
 Qy 721 TATCTCATGTTAAACGTAATAACATACATATCAATTAACGGGATCATGCC 780  
 Db 721 TATCTCATGTTAAACGTAATAACATACATATCAATTAACGGGATCATGCC 780  
 Qy 781 AGAGATTGTTGGGGAAAGTGTATCACGGTTTCAAGTTAATCCAAAAGTT 840  
 Db 781 AGAGATTGTTGGGGAAAGTGTATCACGGTTTCAAGTTAATCCAAAAGTT 840  
 Qy 841 ATTCTCTGAAAAATCAAAATAGTTTCATGTTTCATGTTTCATGTTTAAAA 900  
 Db 841 ATTCTCTGAAAAATCAAAATAGTTTCATGTTTCATGTTTCATGTTTAAAA 900  
 Qy 901 CCACGTGAAATGTTAAATGGACTGTGCAATTCTGCATATACTAAATTA 960  
 Db 901 CCACGTGAAATGTTAAATGGACTGTGCAATTCTGCATATACTAAATTA 960  
 Qy 961 TTAAAGTCATCAGTATCAACATCTTTACACTAAAAGCC 1004  
 Db 961 TTAAAGTCATCAGTATCAACATCTTTACACTAAAAGCC 1004

RESULT 3  
 US-10-363-4 08-1/c  
 / Sequence 1, Application US/10363408  
 / Publication No. US20040077042A1  
 ; GENERAL INFORMATION:  
 / APPLICANT: XIA, Jiahui  
 / TITLE OF INVENTION: A Cell Line Expressing Mutated Human Tissue-Type Plasminogen Activator, Its Construction Strategy and Method of Expression  
 / TITLE OF INVENTION: Activator, Its Construction Strategy and Method of Expression  
 / FILE REFERENCE: 4596-0103P  
 / CURRENT APPLICATION NUMBER: US/10/363,408  
 / CURRENT FILING DATE: 2003-10-17  
 / PRIOR APPLICATION NUMBER: PCT/CN01/00127  
 / PRIOR FILING DATE: 2001-02-16  
 / OTHER INFORMATION: target gene vector  
 / SEQ ID NO: 1  
 / LENGTH: 11162  
 / TYPE: DNA  
 / ORGANISM: Homo sapiens  
 / FEATURE:  
 / NAME/KEY: misc\_feature  
 / LOCATION: (1)..(1162)  
 / OTHER INFORMATION: target gene vector

US-10-363-408-1

Query Match 52.2%; Score 524.4; DB 17; Length 11162;  
 Best Local Similarity 98.9%; Pred. No. 2.2e-127;  
 Matches 528; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 471 TTACAGCGATGTGCAATTGTAAGCACATTGGGCCACTATCAAATGCTG 530  
 Db 6578 TTACAGCAATGTGCAATTGTAAGCACATTGGGCCACTATCAAATGCTG 6519  
 Qy 531 TGTAACCCAGATGTGCAATTGTAAGCACATTGGGCCACTATCAAATGCTG 590  
 Db 6518 TGTAACCCAGATGTGCAATTGTAAGCACATTGGGCCACTATCAAATGCTG 6459  
 Qy 591 CACATCTCTGCTTGTCAAGAACATCAGCAAGCTTACATTGAAAGATCT 650  
 Db 9224 CACATCTCTGCTTGTCAAGAACATCAGCAAGCTTACATTGAAAGATCT 9165  
 Qy 651 ACGCATAGACTTTAACCTAACCAATGTTTCAAGTAACTGAAATGTTGTAATT 710  
 Db 9164 ACGGATAGACTTTAACCTAACCAATGTTTCAAGTAACTGAAATGTTGTAATT 9105  
 Qy 711 CCCTGAGTGTATTCTACTGTTAACGGTAAATGTTGAGGAAATGTTGAGG 770  
 Db 9104 CCCTGAGTGTATTCTACTGTTAACGGTAAATGTTGAGGAAATGTTGAGG 9045  
 Qy 771 GATCATTTGCAAGAGTTGTTGGGGAAATGTTGACACGGTTCAATTGAAATTAAT 830  
 Db 9044 GATCCTTGGCAAGAGTTGTTGGGGAAATGTTGACACGGTTCAATTGAAATTAAT 9985

RESULT 4  
 US-10-362-916-1/c  
 / Sequence 1, Application US/10362916  
 / Publication No. US20040038919A1  
 ; GENERAL INFORMATION:  
 / APPLICANT: XIA, Jiahui  
 / TITLE OF INVENTION: A No. US20040038919A1 Gene Therapy Agent For Hemophilia B And  
 / TITLE OF INVENTION: Method  
 / FILE REFERENCE: 4596-0102P  
 / CURRENT APPLICATION NUMBER: US/10/362,916  
 / CURRENT FILING DATE: 2003-06-30  
 / NUMBER OF SEQ ID NOS: 3  
 / SOFTWARE: Patentin version 3.2  
 / SEQ ID NO: 1  
 / LENGTH: 13928  
 / TYPE: DNA  
 / ORGANISM: Homo sapiens  
 US-10-362-916-1

Query Match 52.2%; Score 524.4; DB 13; Length 13928;  
 Best Local Similarity 98.9%; Pred. No. 2.6e-123;  
 Matches 528; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 471 TTACAGCGATGTGCAATTGTAAGCACATTGGGCCACTATCAAATGCTG 530  
 Db 9344 TTTCACCAATGTGCAATTGTAAGCACATTGGGCCACTATCAAATGCTG 9285  
 Qy 531 TGTAACCCAGATGTGCAATTGTAAGCACATTGGGCCACTATCAAATGCTG 590  
 Db 9284 TGTAACCCAGATGTGCAATTGTAAGCACATTGGGCCACTATCAAATGCTG 9225  
 Qy 591 CACATCTCTGCTTGTCAAGAACATCAGCAAGCTTACATTGAAAGATCT 650  
 Db 9224 CACATCTCTGCTTGTCAAGAACATCAGCAAGCTTACATTGAAAGATCT 9165  
 Qy 651 ACGCATAGACTTTAACCTAACCAATGTTTCAAGTAACTGAAATGTTGTAATT 710  
 Db 9164 ACGGATAGACTTTAACCTAACCAATGTTTCAAGTAACTGAAATGTTGTAATT 9105  
 Qy 711 CCCTGAGTGTATTCTACTGTTAACGGTAAATGTTGAGGAAATGTTGAGG 770  
 Db 9104 CCCTGAGTGTATTCTACTGTTAACGGTAAATGTTGAGGAAATGTTGAGG 9045  
 Qy 771 GATCATTTGCAAGAGTTGTTGGGGAAATGTTGACACGGTTCAATTGAAATTAAT 830  
 Db 9044 GATCCTTGGCAAGAGTTGTTGGGGAAATGTTGACACGGTTCAATTGAAATTAAT 9985

RESULT 5  
 US-10-029-386-1510/c  
 ; Sequence 1510, Application US/10029386  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Penn, Sharron G.  
 ; APPLICANT: Rank, David R.  
 ; APPLICANT: Hanel, David K.  
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GENE EXPRESSION ANALYSIS TWO  
 ; FILE REFERENCE: AEONICA-X-2  
 ; CURRENT APPLICATION NUMBER: US/10/029, 386  
 ; CURRENT FILING DATE: 2001-12-20  
 ; NUMBER OF SEQ ID NOS: 34288  
 ; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1  
 SEQ ID NO: 8714  
 LENGTH: 524  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 OTHER INFORMATION: MAP TO AF254983.1  
 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.95  
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1  
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.81  
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.95  
 OTHER INFORMATION: EST HUMAN HIT: AI138404.1, EVALUATION 0.00e+00  
 OTHER INFORMATION: NT HIT: AU163201.2, EVALUATION 0.00e+00  
 OTHER INFORMATION: SWISSPROT HIT: P55200, EVALUATION 9.00e-04  
 US-10-029-386-1510

Query Match 32.9%; Score 330.4; DB 15; Length 524;  
 Best Local Similarity 98.2%; Pred. No. 2.2e-74; Indels 6; Mismatches 0; Gaps 0;

Qy 471 TTACAGCGATGCGATTGTAACTTGGCCACTTCAATGCTGTGAAGAGAAA 530  
 Db 185 TTTCAGGGATGTTGCAATTGTAAAGCACTTTGGGCCACTTCAATGCTGTGAAGAGAAA 244  
 Qy 531 TGTACCCAGATGTATCATTATCCCTGTGTGCAGGGCGCCCTTCAGATTTCAGT 590  
 Db 245 TGTACCCAGATGTATCATTATCCCTGTGTGCAGGGCGACCTTCAGATTTCAGT 304  
 Qy 591 CACATHTTCTGCCTGCTTGCCAGACATGGCAAGTGTAAATGTTACT 650  
 Db 305 CACATTCCTGCTTGCCAGACATGGCAAGTGTAAATGTTACT 364  
 Qy 651 ACGCATAGACTTTAACACTCAACCATGTTACTGAAATAACAAATGTTGTAATT 710  
 Db 365 ACGCATAGACTTTAACACTCAACCATGTTACTGAAATAACAAATGTTGTAATT 424  
 Qy 711 CCCTGATGTGTATTCTACTGTATTAAAGTTAATCATATACTAAATCTGAGG 770  
 Db 425 CCCTGATGTGTATTCTACTGTATTAAAGTTAATCATATACTAAATCTGAGG 484

RESULT 6  
 US-10-029-386-8714  
 ; Sequence 8714, Application US/10029386  
 ; Publication No. US20030194704A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Penn, Sharron G.  
 ; APPLICANT: Rank, David R.  
 ; APPLICANT: Hanel, David K.  
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GENE EXPRESSION ANALYSIS TWO  
 ; FILE REFERENCE: AEONICA-X-2  
 ; CURRENT APPLICATION NUMBER: US/10/029, 386  
 ; CURRENT FILING DATE: 2001-12-20  
 ; NUMBER OF SEQ ID NOS: 34288

Query Match 32.9%; Score 330.4; DB 15; Length 524;  
 Best Local Similarity 98.2%; Pred. No. 2.2e-74; Indels 6; Mismatches 0; Gaps 0;

Qy 471 TTACAGCGATGCGATTGTAACTTGGCCACTTCAATGCTGTGAAGAGAAA 530  
 Db 185 TTTCAGGGATGTTGCAATTGTAAAGCACTTTGGGCCACTTCAATGCTGTGAAGAGAAA 244  
 Qy 531 TGTACCCAGATGTATCATTATCCCTGTGTGCAGGGCGCCCTTCAGATTTCAGT 590  
 Db 245 TGTACCCAGATGTATCATTATCCCTGTGTGCAGGGCGACCTTCAGATTTCAGT 304  
 Qy 591 CACATHTTCTGCCTGCTTGCCAGACATGGCAAGTGTAAATGTTACT 650  
 Db 305 CACATTCCTGCTTGCCAGACATGGCAAGTGTAAATGTTACT 364  
 Qy 651 ACGCATAGACTTTAACACTCAACCATGTTACTGAAATAACAAATGTTGTAATT 710  
 Db 365 ACGCATAGACTTTAACACTCAACCATGTTACTGAAATAACAAATGTTGTAATT 424  
 Qy 711 CCCTGATGTGTATTCTACTGTATTAAAGTTAATCATATACTAAATCTGAGG 770  
 Db 425 CCCTGATGTGTATTCTACTGTATTAAAGTTAATCATATACTAAATCTGAGG 484

RESULT 7  
 US-10-029-386-18339/c  
 ; Sequence 18339, Application US/10029386  
 ; Publication No. US20030194704A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Penn, Sharron G.  
 ; APPLICANT: Rank, David R.  
 ; APPLICANT: Hanel, David K.  
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GENE EXPRESSION ANALYSIS TWO  
 ; FILE REFERENCE: AEONICA-X-2  
 ; CURRENT APPLICATION NUMBER: US/10/029, 386  
 ; CURRENT FILING DATE: 2001-12-20  
 ; NUMBER OF SEQ ID NOS: 34288

Query Match 32.9%; Score 330.4; DB 15; Length 524;  
 Best Local Similarity 98.2%; Pred. No. 2.2e-74; Indels 6; Mismatches 0; Gaps 0;

Qy 471 TTACAGCGATGCGATTGTAACTTGGCCACTTCAATGCTGTGAAGAGAAA 530  
 Db 340 TTTCAGGGATGTTGCAATTGTAAAGCACTTTGGGCCACTTCAATGCTGTGAAGAGAAA 281  
 Qy 531 TGTACCCAGATGTATCATTATCCCTGTGTGCAGGGCGCCCTTCAGATTTCAGT 590  
 Db 280 TGTACCCAGATGTATCATTATCCCTGTGTGCAGGGCGCCCTTCAGATTTCAGT 221  
 Qy 591 CACATCTCTGCTTGTGCAAGACATGGCAAGCTTCAAGATGTAAATGTTACT 650  
 Db 220 CACATTCCTGCTTGTGCAAGACATGGCAAGCTTCAAGATGTAAATGTTACT 161  
 Qy 651 ACGCATAGACTTTAACACTCAACCATGTTACTGAAATAACAAATGTTGTAATT 710  
 Db 160 ACGCATAGACTTTAACACTCAACCATGTTACTGAAATAACAAATGTTGTAATT 101  
 Qy 711 CCCTGATGTGTATTCTACTGTATTAAAGTTAATCATATACTAAATCTGAGG 770  
 Db 100 CCCTGATGTGTATTCTACTGTATTAAAGTTAATCATATACTAAATCTGAGG 41  
 Qy 771 GATCATGGCAGAGATTGTGGGAGGAATGTATCAA 810  
 Db 40 GATCATGGCAGAGATTGTGGGAGGAATGTATCAA 524

SOFTWARE: Annotmax Sequence Listing Engine vers. 1.1  
SEQ ID NO: 18339  
LENGTH: 345  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AL049849.1  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.9  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.7  
OTHER INFORMATION: EXPRESSED IN PELVIS, SIGNAL = 1.5  
OTHER INFORMATION: NT HIT: 9116165974, EVALUE 0.00e+00  
OTHER INFORMATION: EST HUMAN HIT: BE890756.1, EVALUE 7.00e-65  
OTHER INFORMATION: SWISSPROT HIT: Q13072, EVALUE 2.00e-06  
US-10-029-386-18339

Query Match Score 164.6; DB 15; Length 345;  
Best Local Similarity 97.7%; Pred. No. 6.7e-32; Indels 0; Gaps 0;  
Matches 167; Conservative 0; Mismatches 4;

QY 215 GGTTTTCTGGCATGTCGCCAACCTGCTCCANCCAGCGCTGATGAGGAGTC 274  
Db 282 GGTTTTCTGGCATGTCGCCAACCTGCTCCAGCCAACTGTGAGGAGTC 223

QY 275 TGTGTTGAGCTGGAGTTGAGCCTGAAGGGAGCTGTGCTCATCTCTGG 334  
Db 222 TGTGTTGAGCTGGAGTTGAGCCTGAAGGGAGCTGTGCTCATCTCTGG 163

QY 335 TTGTGGAGCCACGGTGATGAGAAGGGAGCTAACAGGGCAATAGGAGG 385  
Db 162 TTGTGGAGCCACGGTGATGAGAAGGGAGCTAACAGGGCAATAGGAGG 112

RESULT 8  
US-10-029-386-4583/c  
Sequence 4533, Application US/10029386  
Publication No. US20030194704A1  
CURRENT APPLICATION NUMBER: US/10/029,386  
CURRENT FILING DATE: 2001-12-20  
NUMBER OF SEQ ID NOS: 34288  
SOFTWARE: Annotmax Sequence Listing Engine vers. 1.1  
SEQ ID NO: 4583  
LENGTH: 538  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AL049849.1  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.7  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5  
OTHER INFORMATION: NT HIT: S114763849, EVALUE 4.00e-06  
OTHER INFORMATION: EST HUMAN HIT: BE890756.1, EVALUE 1.00e-64  
OTHER INFORMATION: EST\_HUMAN HIT: BE890756.1, EVALUE 1.00e-64  
US-10-029-386-4583

Query Match Score 164.6; DB 15; Length 538;  
Best Local Similarity 97.7%; Pred. No. 9e-32; Indels 0; Gaps 0;  
Matches 167; Conservative 0; Mismatches 4;

QY 215 GGTTTTCTGGCATGTCGCCAACCTGCTCCAGCCAACTGTGAGGAGTC 274

RESULT 9  
US-10-029-386-22449  
Sequence 22449, Application US/10029386  
Publication No. US20030194704A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharron G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanze, David K.  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
TITLE OF INVENTION: EXPRESSION ANALYSIS TWO  
FILE REFERENCE: AEOMICA-X-2  
CURRENT APPLICATION NUMBER: US/10/029,386  
CURRENT FILING DATE: 2001-12-20  
NUMBER OF SEQ ID NOS: 34288  
SOFTWARE: Annotmax Sequence Listing Engine vers. 1.1  
SEQ ID NO: 22449  
LENGTH: 172  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.95  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.81  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.95  
OTHER INFORMATION: SWISSPROT HIT: PS5200, EVALUE 2.00e-04  
OTHER INFORMATION: EST\_HUMAN HIT: AL138404.1, EVALUE 6.00e-82  
OTHER INFORMATION: NT HIT: AL163201.2, EVALUE 3.00e-92  
US-10-029-386-22449

Query Match Score 164.2; DB 15; Length 172;  
Best Local Similarity 98.2%; Pred. No. 5.5e-32;  
Matches 166; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 471 TTACAGGATGNGCATTTGTAAGGACTTTGGCCACTATCAATGCTGTGAAGAGAAA 530  
Db 4 TTACAGGATGNGCATTTGTAAGGACTTTGGCCACTATCAATGCTGTGAAGAGAAA 63

QY 531 TGTACCCAGATGTATCATTATCCTGTCGAGGATTCAGT 590  
Db 64 TGTACCCAGATGTATCATTATCCTGTCGAGGATTCAGT 123

QY 591 CACATCTTCCCTGCTTGTCCAGAAACATGGCCAGCTCTGTGAAGAGAT 639  
Db 124 CACATCTTCCCTGCTTGTCCAGAAACATGGCCAGCTCTGTGAAGAGAT 172

RESULT 10  
US-10-029-386-15211/c  
Sequence 15211, Application US/10029386  
Publication No. US20030194704A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharron G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanze, David K.  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
TITLE OF INVENTION: EXPRESSION ANALYSIS TWO  
FILE REFERENCE: AEOMICA-X-2  
CURRENT APPLICATION NUMBER: US/10/029,386  
CURRENT FILING DATE: 2001-12-20  
NUMBER OF SEQ ID NOS: 34288



PRIOR APPLICATION NUMBER: US 60/198,676  
 PRIOR FILING DATE: 2000-04-20  
 PRIOR APPLICATION NUMBER: US 60/193,483  
 PRIOR APPLICATION NUMBER: US 60/185,218  
 PRIOR FILING DATE: 2000-03-29  
 PRIOR APPLICATION NUMBER: US 60/185,218  
 PRIOR FILING DATE: 2000-02-24  
 PRIOR APPLICATION NUMBER: US 60/167,363  
 PRIOR FILING DATE: 1999-11-23  
 PRIOR APPLICATION NUMBER: US 60/156,358  
 PRIOR FILING DATE: 1999-09-28  
 PRIOR APPLICATION NUMBER: US 60/146,002  
 PRIOR FILING DATE: 1999-08-09  
 SOFTWARE: FastSEQ for Windows Version 4.0  
 SEQ ID NO: 128026  
 LENGTH: 407  
 TYPE: DNA  
 ORGANISM: Human  
 US-10-027-632-128026

Query Match 8.8%; Score 88.4; DB 13; Length 407;  
 Best Local Similarity 98.9%; Pred. No. 2.8e-12;  
 Matches 89; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 471 TTACAGCGATGCGATTGTAAGCACTTGGACCACTATAATGGCTGTGAGAGAA 530  
 Db 90 TTTCAGCGATGCGATTGTAAGCACTTGGACCACTATAATGGCTGTGAGAGAA 31

Qy 531 TGTACCCAGATGTATCATATCCTGTGCT 560  
 Db 30 .TGTACCCAGATGTATCATATCCTGTGCT 1

RESULT 14  
 US-10-027-632-128027/C  
 Sequence 128027, Application US/10027632

Publication No. US20020198371A1  
 GENERAL INFORMATION:

APPLICANT: Wang, David G.  
 TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
 FILE REFERENCE: 108827-129  
 CURRENT APPLICATION NUMBER: US/10/027,632  
 CURRENT FILING DATE: 2002-04-30  
 PRIOR APPLICATION NUMBER: US 60/218,006  
 PRIOR FILING DATE: 2000-07-12  
 PRIOR APPLICATION NUMBER: US 60/198,676  
 PRIOR FILING DATE: 2000-04-20  
 PRIOR APPLICATION NUMBER: US 60/193,483  
 PRIOR FILING DATE: 2000-03-29  
 PRIOR APPLICATION NUMBER: US 60/185,218  
 PRIOR FILING DATE: 2000-02-24  
 PRIOR APPLICATION NUMBER: US 60/167,363  
 PRIOR FILING DATE: 1999-11-23  
 PRIOR APPLICATION NUMBER: US 60/156,358  
 PRIOR FILING DATE: 1999-09-28  
 PRIOR APPLICATION NUMBER: US 60/146,002  
 PRIOR FILING DATE: 1999-08-09  
 NUMBER OF SEQ ID NOS: 325720  
 SOFTWARE: FastSEQ for Windows Version 4.0  
 SEQ ID NO: 128028  
 LENGTH: 407  
 TYPE: DNA  
 ORGANISM: Human  
 US-10-027-632-128028

Query Match 8.8%; Score 88.4; DB 13; Length 407;  
 Best Local Similarity 98.9%; Pred. No. 2.8e-12;  
 Matches 89; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 471 TTACAGCGATGCGATTGTAAGCACTTGGACCACTATAATGGCTGTGAGAGAA 530  
 Db 90 TTTCAGCGATGCGATTGTAAGCACTTGGACCACTATAATGGCTGTGAGAGAA 31

Qy 531 TGTACCCAGATGTATCATATCCTGTGCT 560  
 Db 30 .TGTACCCAGATGTATCATATCCTGTGCT 1

Search completed: July 9, 2004, 12:21:50  
 Job time : 697 secs

Query Match 8.8%; Score 88.4; DB 13; Length 407;  
 Best Local Similarity 98.9%; Pred. No. 2.8e-12;  
 Matches 89; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 471 TTACAGCGATGCGATTGTAAGCACTTGGACCACTATAATGGCTGTGAGAGAA 530  
 Db 90 TTTCAGCGATGCGATTGTAAGCACTTGGACCACTATAATGGCTGTGAGAGAA 31

This Page Blank (uspto)

This Page Blank (uspto)

GenCore version 5.1.6									
Copyright (c) 1993 - 2004 Compugen Ltd.									
OM nucleic - nucleic search, using sw model									
Run on:	July 9, 2004, 03:21:18 ; Search time 4192 Seconds (without alignments)								
	7152.106 Million cell updates/sec								
Title:	US-10-081-108-1								
Perfect score:	1004								
Sequence:	CGCCGAATTAGGGCTCCGG.....ATCTTTAGACTAAAAAGCC 1004								
Scoring table:	IDENTITY NUC								
	Gapo 10.0 , Gapext 1.0								
Searched:	2751389 seqs, 1493109026 residues								
Total number of hits satisfying chosen parameters:	55026578								
Minimum DB seq length: 0									
Maximum DB seq length: 20000000000									
Post-processing: Minimum Match 0% Maximum Match 100% Listing First 45 summaries									
Database :	EST:*								
	1: em_estba:*								
	2: em_esthum:*								
	3: em_estrin:*								
	4: em_estmu:*								
	5: em_estcov:*								
	6: em_estpl:*								
	7: em_estro:*								
	8: em_htc:*								
	9: gb_est1:*								
	10: gb_est2:*								
	11: gb_estc:*								
	12: gb_est3:*								
	13: gb_est4:*								
	14: gb_est5:*								
	15: em_estfun:*								
	16: em_estbom:*								
	17: em_gss_hum:*								
	18: em_gss_irv:*								
	19: em_gss_pdn:*								
	20: em_gss_yrt:*								
	21: em_gss_fun:*								
	22: em_gss_man:*								
	23: em_gss_mus:*								
	24: em_gss_pro:*								
	25: em_gss_rdc:*								
	26: em_gss_plg:*								
	27: em_gss_vri:*								
	28: gb_gssi:*								
	29: gb_gss2:*								
RESULTS									
1	BC017024	LOCUS	Homo sapiens	cdNA clone IMAGE:3916415, containing frame-shift errors	1	DEFINITION	Linear mRNA	18-JUN-2000	
	BC017024	ACCESSION				VERSION			
	BC017024	KEYWORDS				HTC.			
	BC017024	ORGANISM	Homo sapiens (human)						
	BC017024	SOURCE	Eukaryota; Bacteria; Chordata; Craniata; Catarrhini; Hominidae; Homo.						
	BC017024	AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.						
	BC017024	REFERENCE	Krausser, R.D., Strainsberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Schubler, G.D., Altrichul, S.F., Zeeberg, B.R., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Sores, M.B., Bonaldo, M.P., Casavant, T.L., Schuetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S., Loqueline, N.A., Peters, G.J., Abramson, R.D., Mulyahy, S.J., Bosak, S.A., McBryan, P.J., McFernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.W., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Murray, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahney, J., Helton, E., Kertesz, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,						
	BC017024	Description	Homo sapiens						
	BC017024		1						
	BC017024		2						
	BC017024		3						
	BC017024		4						
	BC017024		5						
	BC017024		6						
	BC017024		7						
	BC017024		8						
	BC017024		9						
	BC017024		10						
	BC017024		11						
	BC017024		12						
	BC017024		13						
	BC017024		14						
	BC017024		15						
	BC017024		16						
	BC017024		17						
	BC017024		18						
	BC017024		19						
	BC017024		20						
	BC017024		21						
	BC017024		22						
	BC017024		23						
	BC017024		24						
	BC017024		25						
	BC017024		26						
	BC017024		27						
	BC017024		28						
	BC017024		29						
	BC017024		30						
	BC017024		31						
	BC017024		32						
	BC017024		33						
	BC017024		34						
	BC017024		35						
	BC017024		36						
	BC017024		37						
	BC017024		38						
	BC017024		39						
	BC017024		40						
	BC017024		41						
	BC017024		42						
	BC017024		43						
	BC017024		44						
	BC017024		45						
	BC017024		46						
	BC017024		47						
	BC017024		48						
	BC017024		49						
	BC017024		50						
	BC017024		51						
	BC017024		52						
	BC017024		53						
	BC017024		54						
	BC017024		55						
	BC017024		56						
	BC017024		57						
	BC017024		58						
	BC017024		59						
	BC017024		60						
	BC017024		61						
	BC017024		62						
	BC017024		63						
	BC017024		64						
	BC017024		65						
	BC017024		66						
	BC017024		67						
	BC017024		68						
	BC017024		69						
	BC017024		70						
	BC017024		71						
	BC017024		72						
	BC017024		73						
	BC017024		74						
	BC017024		75						
	BC017024		76						
	BC017024		77						
	BC017024		78						
	BC017024		79						
	BC017024		80						
	BC017024		81						
	BC017024		82						
	BC017024		83						
	BC017024		84						
	BC017024		85						
	BC017024		86						
	BC017024		87						
	BC017024		88						
	BC017024		89						
	BC017024		90						
	BC017024		91						
	BC017024		92						
	BC017024		93						
	BC017024		94						
	BC017024		95						
	BC017024		96						
	BC017024		97						
	BC017024		98						
	BC017024		99						
	BC017024		100						
	BC017024		101						
	BC017024		102						
	BC017024		103						
	BC017024		104						
	BC017024		105						
	BC017024		106						
	BC017024		107						
	BC017024		108						
	BC017024		109						
	BC017024		110						
	BC017024		111						
	BC017024		112						
	BC017024		113						
	BC017024		114						
	BC017024		115						
	BC017024		116						
	BC017024		117						
	BC017024		118						
	BC017024		119						
	BC017024		120						
	BC017024		121						
	BC017024		122						
	BC017024		123						
	BC017024		124						
	BC017024		125						
	BC017024		126						
	BC017024		127						
	BC017024		128						
	BC017024		129						
	BC017024		130						
	BC017024		131						
	BC017024		132						
	BC017024		133						
	BC017024		134						
	BC017024		135						
	BC017024		136						
	BC017024		137						
	BC017024		138						
	BC017024		139						
	BC017024		140						
	BC017024		141						
	BC017024		142						
	BC017024		143						
	BC017024		144						
	BC017024		145						
	BC017024		146						
	BC017024		147						
	BC017024		148						
	BC017024		149						
	BC017024		150						
	BC017024		151						
	BC017024		152						
	BC017024								

325	AAGGGGACAGCTGCTTCACTTGCTTCAATTTGAGGTGAGGAGCG	384
Dickson, M.C., Rodriguez, A.C., Grinwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalius, D.E., Schnarch, A., Schein, J.E., Jones, S.J. and Marra, M.A.	361 GCAGCTCACAGGACAATTAGGAG 385 GCAGTCACAGGACAATTAGGAGTAACCTGAGGTAACAGTGCAGGTTCAAA	400 444 Generation and initial analyses of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)	401 GTGTAAGCCAGGATGTCGATCTCCGATCCTGATCCTGGCTTGGCTTCAA	460
MEDLINE 22388257	445 GTGTAAGCCAGGATGTCGATCTCCGATCCTGATCCTGGCTTGGCTTCAA	504
REFERENCE 2 (bases 1 to 2037)	461 AGTGGCGAGATA 505 AGTGGCGAGATAAGGGTTATAGAGATTTCTCCCTTACCGATGTCG	483 564
AUTHORS Straubhaar, R.	Db Direct Submission	
JOURNAL Submitted (05-NOV-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	Db 484 CATTGTGAGACTTGAGCTTGAGGCCATCAATGCTGAAGAGAAATGACCGAGTGT 565 CATTGTGAGACTTGAGGCCATCAATGCTGAAGAGAAATGACCGAGTGT	543 624
REMARK NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>	Db 544 ATCATATTCTGGCTGAGGAGGGCTCCCTCAGATTCTGACATCTTCTCC 625 ATCATATTCTGGCTGAGGAGGGCTCCCTCAGATTCTGACATCTTCTCC	603 684
COMMENT Contact: MGC help desk	Qy 544 ATCATATTCTGGCTGAGGAGGGCTCCCTCAGATTCTGACATCTTCTCC 625 ATCATATTCTGGCTGAGGAGGGCTCCCTCAGATTCTGACATCTTCTCC	
Email: cgpbs-r@mail.nih.gov	Qy 604 TTGGTCAGAACACATTGACAGGTGAGTTGAGCTAGCATAGACTTT 685 TTGGTCAGAACATTTGAGCTGAGCTGAGCTGAGTTGAGCTAGCATAGACTTT	
Tissue Procurement: ATCC/DCTP/DTP	Db 685 TTGGTCAGAACATTTGAGCTGAGCTGAGCTGAGTTGAGCTAGCATAGACTTT	663
CDNA Library Preparation: Life Technologies, Inc.	Qy 685 TTGGTCAGAACATTTGAGCTGAGCTGAGCTGAGTTGAGCTAGCATAGACTTT	744
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)	Db 685 TTGGTCAGAACATTTGAGCTGAGCTGAGCTGAGTTGAGCTAGCATAGACTTT	
Gene Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305	Qy 685 TTGGTCAGAACATTTGAGCTGAGCTGAGCTGAGTTGAGCTAGCATAGACTTT	
Web site: <a href="http://www-shgc.stanford.edu">http://www-shgc.stanford.edu</a>	Db 723 TAAACTTCACCAATGATTACTGAAATAACAAATGTTAAATCCCTGAGTGTAT 745 TAAACTTCACCAATGATTACTGAAATAACAAATGTTAAATCCCTGAGTGTAT	
Contact: (Dickson, Mark) mdc@paxil.stanford.edu	Qy 723 TAAACTTCACCAATGATTACTGAAATAACAAATGTTAAATCCCTGAGTGTAT Db 745 TAAACTTCACCAATGATTACTGAAATAACAAATGTTAAATCCCTGAGTGTAT	
Dickson, M., Schmutz, J., Grinwood, J., Rodriguez, A., and Myers, R. M.	Db 745 TAAACTTCACCAATGATTACTGAAATAACAAATGTTAAATCCCTGAGTGTAT 804	
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium URL at: <a href="http://image.lnl.gov">http://image.lnl.gov</a>	Qy 724 TCTACTTGTATTAAAGGTTAAATACATAATCATAAATCTGAGGGATCATGCCAGA 805 TCTACTTGTATTAAAGGTTAAATACATAATCATAAATCTGAGGGATCATGCCAGA	783 864
Series: IRAK plate: 22 Row: o Column: 15	Db 805 TCTACTTGTATTAAAGGTTAAATACATAATCATAAATCTGAGGGATCATGCCAGA	
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4557346	Qy 784 GATTGTTGGAGGGAAATGTTAACACCGTTCTATTGAAATTAAATCCAAAAGTTATT Db 865 GATTGTTGGAGGGAAATGTTAACACCGTTCTATTGAAATTAAATCCAAAAGTTATT	843 924
This clone has the following problem: frame shifted.	Qy 784 GATTGTTGGAGGGAAATGTTAACACCGTTCTATTGAAATTAAATCCAAAAGTTATT Db 865 GATTGTTGGAGGGAAATGTTAACACCGTTCTATTGAAATTAAATCCAAAAGTTATT	
Location/Qualifiers 1..2037	Qy 844 TCCTCAGAAATACTCAAATAAGTTGCTATTTTATCTTAAACATTTTAAACCA Db 925 TCCTCAGAAATACTCAAATAAGTTGCTATTTTATCTTAAACATTTTAAACCA	903 984
/organism="Homo sapiens"	Qy 904 CTGTAAATGTGTAAATAGGGACTGTGAGTATTTGCTACATTAATTTAATTA Db 985 CTGTAAATGTGTAAATAGGGACTGTGAGTATTTGCTACATTAATTTAATTA	963 1044
/mol_type="mRNA"	Qy 985 CTGTAAATGTGTAAATAGGGACTGTGAGTATTTGCTACATTAATTTAATTA	
/db_xref="EAXON:9606"	Db 1045 AAAATGCAATGATGTTACATTTTACACTAAAGGCC	1085
/clone="IMAGE:391415"		
/tissue_type="Skin, melanotic melanoma."		
/cloneLib="NIH MGC_72"		
/lab_host="DH10B"		
/note="vector: PCMV-SPORT6"		
SEQUENCE		
source		
Query Match 92.0%	Score 923.8; DB 11; Length 2037;	
Best Local Similarity 94.4%	Pred. No. 3.6e-220; Mismatches 0; Indels 57; Gaps 2;	
Matches 1002; Conservative		
y 1 CGCCATTAGGGCTCCCGTATTCGGTACCGCTGAGCTGGCTTAGAGACC 60	RESULT 2 CA423957/c	732 bp mRNA linear EST 07-NOV-2002
y 25 CGCCATTAGGGCTCCCGTATTCGGTACCGCTGAGCTGGCTTAGAGACC 84	LOCUS CA423957	UTR-H-PEI-bed-p-04-0-UTI s1 NCI CGAP FET Homo sapiens cDNA clone
o 61 AGGAAAGCCGAACTGGCTGAGCTGAGCTGAGCTGGCTGAGCTGGT 120	DEFINITION CA423957	UTR-H-PEI-bed-p-04-0-UTI 3' mRNA sequence.
o 85 AGGAAAGCCGAACTGGCTGAGCTGAGCTGGCTGAGCTGGT 144	VERSION CA423957.1	1 (bases 1 to 732)
o 121 GTGGTGGCAACAGAGATGACAGCTGGAGTGTAGGAGGCTGAGCTGGT 180	KEYWORDS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
o 145 GTGGTGGCAACAGAGATGACAGCTGGAGTGTAGGAGGCTGAGCTGGT 204	ORGANISM Homo sapiens (human)	Tumor Gene Index
o 181 AGTGGGGCTGGAGGAGTAAATGCGCCGAGCGTTCTGGCATGTCGCCAGC 240	REFERENCE Unpublished (1997)	
o 205 AGTGGGGCTGGAGGAGTAAATGCGCCGAGCGTTCTGGCATGTCGCCAGC 264	AUTHORS Contact: Robert Strausberg, Ph.D.	
b 241 TGCTCCAAACGAGCTGAGAGGAGTCCCTGAGCTGAGGTTGGAGCTG 300	Tissue Preparation: James Martin	
b 265 TGCTCCAAACGAGCTGAGAGGAGTCCCTGAGCTGAGGTTGGAGCTG 324	Email: cgaps-r@mail.nih.gov	
y 301 AAGGGGACAGCTGCTTCACTCTGAGTTGGCACCCAGCTGATGCCAGC 360	CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa	

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares bento-soares@uiowa.edu  
 The following repetitive elements were found in this cDNA sequence: 391-493, >ALU 493-590, >ALU  
 Seq\_Primer: M13 FORWARD  
 POLYA=Yes  
 FEATURES  
 source 1.  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9006"  
 /clone="JUL-H-FE1-ped-P-04-0-UI"  
 /tissue\_type="Cell lines"  
 /dev\_stage="Adult"  
 /lab\_host="DH10B (Life Technologies)"  
 /clone\_lib="NCI CGAP\_FPL"  
 /note="Organ: Chondrosarcoma; Vector: pTR3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP\_FPL is a normalized cDNA library derived from a pool of mRNA obtained from 3 cell lines from Grade II chondrosarcoma tissues. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6: 791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pTR3-Pac vector. The oligonucleotide used to prime the synthesis of first strand cDNA contains a library tag (dtr18 tail). The sequence tag for this library is CGCTAACGGAC. The cell lines were provided by Dr James Martin from the University of Iowa.  
 TAG TISSUE=Human grade 2 chondrosarcoma cell line pool  
 TAG SEQ=CGCTAACGGAC

## ORIGIN

Query Match 44.9%; Score 450.8; DB 14; Length 732;  
 Best Local Similarity 81.5%; Pred. No. 8.1e-102; Gaps 2;  
 Matches 595; Conservative 0; Mismatches 27; Indels 108;

QY 243 CTCCAGGCCAGGCTATGAGGAGACTCCCTGGTAGCTGGAGTTGGCCCAA 302  
 Db 732 CTCCAGGCCAGGCTATGAGGAGACTCCCTGGTAGCTGGAGTTGGCCCAA 673  
 QY 303 GACGGCAAGCTCTGTCTCATCTCTGAGGTGTGGAGCACGGATGAGACGC 362  
 Db 672 GACGGCAAGCTCTGTCTCATCTCTGAGGTGTGGAGCACGGATGAGACGC 613  
 QY 363 AGCTCAAGGAAATAAGGAGATGGAGTTCATCTGTGAGGTGGCCGGCTATCTGA 422  
 Db 612 AGCTCAAGGAAATAAGGAGATGGAGCAGGCACTCTCGATGTTGGCCGGCTATCTGA 553  
 QY 423 TCTCCCTGACCTT-GTGATCGCCGCTTCCCAA-----461  
 Db 552 GCTCTGGTCAAGTGTGATGCCGCCAGCTCCAAAAGTGTGGATTATAGCG 493  
 QY 462 -----461  
 Db 492 CTGGAGTGCATGGCACGATCTGGCTCACTGGCCATCCGTGGCTCCAGGTCAA 433  
 QY 462 -----461  
 Db 432 GCGATTCTCCGCTCAACCTCCGGAGTAGCTGGATACAGGATGTGCAATTGGAG 373  
 QY 495 CACTTGGAGCCACTATCAAATGCTGTGAGAATGACCCAGATATCATTATCT 554  
 Db 372 CACTTGGAGCCACTATCAAATGCTGTGAGAATGACCCAGATATCATTATCT 313  
 QY 555 TGTGCTGAGGAGCCGCTCTTCAGGATTCACTCCTCCGTCTGAGAATGACCCAGAT 614  
 Db 312 TGTGCTGAGGAGCCGCTCTTCAGGATTCACTCCTCCGTCTGAGAATGACCCAGAT 253

QY 615 CACATGGACCAAGCTCCCTGAAAGATGTAAGTTTACAGCTAGACTTAAACTCAAC 674  
 Db 252 CACATGGACCAAGCTCCCTGAAAGATGTAAGTTTACAGCTAGACTTAAACTCAAC 193  
 QY 675 CAATGTTTACTGAAATAACAACTGTTGAAATTCCTAGTGTATTCTACTGTAT 734  
 Db 192 CAATGTTTACTGAAATAACAACTGTTGAAATTCCTAGTGTATTCTACTGTAT 133  
 QY 735 TAAAGGTTATAATACATAATCATATAATCTGAGGATCATGGCAGAGATTGGGG 794  
 Db 132 TAAAGGTTATAATACATAATCATATAATCTGAGGATCATGGCAGAGTTGGGG 73  
 QY 795 AGGGAAATGTTATCAAAGGTTCATGAAATTAATCCTCAGAAAA 854  
 Db 72 AGGGAAATGTTATCAAAGGTTCATGAAATTAATCCTCAGAAAA 13  
 RESULT 3  
 LOCUS BE890756 699 bp mRNA linear EST 20-OCT-2000  
 DEFINITION 60143191P1 NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE:3916415 5', mRNA sequence.  
 ACCESSION BE890756  
 VERSION BE890756.1 GI:10349397  
 SOURCE EST  
 ORGANISM Homo sapiens (human)  
 KEYWORDS Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 699)  
 AUTHORS NIH MGC <http://mgc.ncbi.nlm.nih.gov/>  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgaps-r@mail.nih.gov  
 Tissue Procurement: ATCC/DCID/DRP  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov/PlatE/IMAGE97.html> row: e column: 24  
 High quality sequence stop: 407.  
 Location/Qualifiers 1..699  
 FEATURES source  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3916415"  
 /tissue\_type="melanotic melanoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_id="NTH MGC\_72"  
 /note="Organ: skin; Vector: PCMV-SPORT6; Site\_1: NotI; Site\_2: Sall; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies."  
 ORIGIN

Query Match 34.8%; Score 349.6; DB 10; Length 699;  
 Best Local Similarity 89.1%; Pred. No. 1.8e-76;  
 Matches 454; Conservative 0; Mismatches 30; Gaps 6;  
 QY 1 CGCCAAATTAGGTCTCCGTTATCCTCCGGTGTAGCTGTGCTGGCTTGGACCC 60  
 Db 4 CGCCAAATTAGGTCTCCGTTATCCTCCGGTGTAGCTGTGCTGGCTTGGACCC 63  
 QY 61 AGGAAAGGGGAGGGCTGGAGGCTGCTGAAACCGTGGCTGAAACCGTGGCTGATCTGGAT 120



ORIGIN									
<i>/clone="KAT00969"</i> <i>/clone_lib="sugano Homo sapiens cDNA library"</i>									
Matches 267; Conservative 0; Mismatches 17; Indels 1; Gaps 1;									
Best Local Similarity 93.7%; Pred. No. 1.4e-50;	Score 245.8; DB 9; Length 300;	Qy 17 CCGTATCTCCGCTGAGCTCTGTTCCGGCTTAGAGGACAGGAAGGGGACT 76	Db 17 CCTGATCTCCGCTGAGCTCTAGTCCGGCTTAGAGGACAGGAAGGGGACT 76	Qy 17 CGGTATCTCCGCTGAGCTCTGTTCCGGCTTAGAGGACAGGAAGGGGACT 76	Db 17 CCTGATCTCCGCTGAGCTCTAGTCCGGCTTAGAGGACAGGAAGGGGACT 76	Qy 17 CGGTATCTCCGCTGAGCTCTGTTCCGGCTTAGAGGACAGGAAGGGGACT 76	Db 17 CCTGATCTCCGCTGAGCTCTAGTCCGGCTTAGAGGACAGGAAGGGGACT 76	Qy 17 CGGTATCTCCGCTGAGCTCTGTTCCGGCTTAGAGGACAGGAAGGGGACT 76	Db 17 CCTGATCTCCGCTGAGCTCTAGTCCGGCTTAGAGGACAGGAAGGGGACT 76
DEFINITION	LOCUS	DEFINITION	LOCUS	DEFINITION	LOCUS	DEFINITION	LOCUS	DEFINITION	LOCUS
3', similar to TR:O1686 AIR.; mRNA sequence.	DEFINITION	3', similar to TR:O1686 AIR.; mRNA sequence.	DEFINITION	3', similar to TR:O1686 AIR.; mRNA sequence.	DEFINITION	3', similar to TR:O1686 AIR.; mRNA sequence.	DEFINITION	3', similar to TR:O1686 AIR.; mRNA sequence.	DEFINITION
SEQUENCE	SEQUENCE	SEQUENCE	SEQUENCE	SEQUENCE	SEQUENCE	SEQUENCE	SEQUENCE	SEQUENCE	SEQUENCE
RESULTS	RESULTS	RESULTS	RESULTS	RESULTS	RESULTS	RESULTS	RESULTS	RESULTS	RESULTS
Qy 197 TAAGATGGGGGAGCTGGAGCTTGAGGGGGCTCTAGAAGTGGCTGGAG 195	Db 196 TAAGATGGGAAGCCGGCTGAGCTTGAGGGGGCTCTAGAAGTGGCTGGAG 196	Qy 197 TAAGATGGGGGAGCTGGAGCTTGAGGGGGCTCTAGAAGTGGCTGGAG 195	Db 196 TAAGATGGGAAGCCGGCTGAGCTTGAGGGGGCTCTAGAAGTGGCTGGAG 196	Qy 197 TAAGATGGGGGAGCTGGAGCTTGAGGGGGCTCTAGAAGTGGCTGGAG 195	Db 196 TAAGATGGGAAGCCGGCTGAGCTTGAGGGGGCTCTAGAAGTGGCTGGAG 196	Qy 197 TAAGATGGGGGAGCTGGAGCTTGAGGGGGCTCTAGAAGTGGCTGGAG 195	Db 196 TAAGATGGGAAGCCGGCTGAGCTTGAGGGGGCTCTAGAAGTGGCTGGAG 196	Qy 197 TAAGATGGGGGAGCTGGAGCTTGAGGGGGCTCTAGAAGTGGCTGGAG 195	Db 196 TAAGATGGGAAGCCGGCTGAGCTTGAGGGGGCTCTAGAAGTGGCTGGAG 196
VERSION	VERSION	VERSION	VERSION	VERSION	VERSION	VERSION	VERSION	VERSION	VERSION
KEYWORDS	KEYWORDS	KEYWORDS	KEYWORDS	KEYWORDS	KEYWORDS	KEYWORDS	KEYWORDS	KEYWORDS	KEYWORDS
SOURCE	SOURCE	SOURCE	SOURCE	SOURCE	SOURCE	SOURCE	SOURCE	SOURCE	SOURCE
ORGANISM	ORGANISM	ORGANISM	ORGANISM	ORGANISM	ORGANISM	ORGANISM	ORGANISM	ORGANISM	ORGANISM
Homo sapiens (human)	Homo sapiens (human)	Homo sapiens (human)	Homo sapiens (human)	Homo sapiens (human)	Homo sapiens (human)	Homo sapiens (human)	Homo sapiens (human)	Homo sapiens (human)	Homo sapiens (human)
COMMENT	COMMENT	COMMENT	COMMENT	COMMENT	COMMENT	COMMENT	COMMENT	COMMENT	COMMENT
JOURNAL	JOURNAL	JOURNAL	JOURNAL	JOURNAL	JOURNAL	JOURNAL	JOURNAL	JOURNAL	JOURNAL
UNPUBLISHED	UNPUBLISHED	UNPUBLISHED	UNPUBLISHED	UNPUBLISHED	UNPUBLISHED	UNPUBLISHED	UNPUBLISHED	UNPUBLISHED	UNPUBLISHED
EST.	EST.	EST.	EST.	EST.	EST.	EST.	EST.	EST.	EST.
HOME PAGE	HOME PAGE	HOME PAGE	HOME PAGE	HOME PAGE	HOME PAGE	HOME PAGE	HOME PAGE	HOME PAGE	HOME PAGE
http://image.llnl.gov	http://image.llnl.gov	http://image.llnl.gov	http://image.llnl.gov	http://image.llnl.gov	http://image.llnl.gov	http://image.llnl.gov	http://image.llnl.gov	http://image.llnl.gov	http://image.llnl.gov
PIPLINE	PIPLINE	PIPLINE	PIPLINE	PIPLINE	PIPLINE	PIPLINE	PIPLINE	PIPLINE	PIPLINE
LIAISON	LIAISON	LIAISON	LIAISON	LIAISON	LIAISON	LIAISON	LIAISON	LIAISON	LIAISON
DATA SOURCE	DATA SOURCE	DATA SOURCE	DATA SOURCE	DATA SOURCE	DATA SOURCE	DATA SOURCE	DATA SOURCE	DATA SOURCE	DATA SOURCE
FEATURES	FEATURES	FEATURES	FEATURES	FEATURES	FEATURES	FEATURES	FEATURES	FEATURES	FEATURES
source	source	source	source	source	source	source	source	source	source
ORIGIN	ORIGIN	ORIGIN	ORIGIN	ORIGIN	ORIGIN	ORIGIN	ORIGIN	ORIGIN	ORIGIN
Query Match Score 15.8%; Best Local Similarity 99.4%; Matches 159; Conservative 1; Indels 0; Gaps 0;	Query Match Score 158.4%; Best Local Similarity 99.4%; Matches 159; Conservative 1; Indels 0; Gaps 0;	Query Match Score 169.4%; Best Local Similarity 99.4%; Matches 160; Conservative 1; Indels 0; Gaps 0;	Query Match Score 16.8%; Best Local Similarity 99.4%; Matches 161; Conservative 1; Indels 0; Gaps 0;	Query Match Score 158.4%; Best Local Similarity 99.4%; Matches 162; Length 940; Indels 0; Gaps 0;	Query Match Score 158.4%; Best Local Similarity 99.4%; Matches 163; Length 945; Indels 0; Gaps 0;	Query Match Score 158.4%; Best Local Similarity 99.4%; Matches 164; Length 945; Indels 0; Gaps 0;	Query Match Score 158.4%; Best Local Similarity 99.4%; Matches 165; Length 945; Indels 0; Gaps 0;	Query Match Score 158.4%; Best Local Similarity 99.4%; Matches 166; Length 945; Indels 0; Gaps 0;	Query Match Score 158.4%; Best Local Similarity 99.4%; Matches 167; Length 945; Indels 0; Gaps 0;



RESULT 10  
 H55516 H55516 160 bp mRNA linear EST 07-DEC-1995  
 LOCUS CHR220455 Chromosome 22 exon Homo sapiens cDNA clone C22\_616\_5,  
 DEFINITION mRNA sequence.

ACCESSION H55516.1 GI:1108382  
 VERSION EST  
 KEYWORDS SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 Hillier,L., Allen,M., Bowles,L., Dubugue,T., Geisel,G., Jost,S.,  
 Kucaba,T., Lacy,M., Le N., Lennon,G., Marra,M., Martin,J.,  
 Moore,B., Schein,E., Shifman,K., Steptoe,M., Tan,F., Theising,B.,  
 White,Y., Wylie,T., Waterston,R. and Wilson,R.  
 WashU-Merck EST Project 1997  
 Unpublished (1997)  
 Contact: Wilson RK  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St.. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available royalty-free through LLNL ; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 Seq. Primer: -28m13 rev2 ET from Amerham.  
 Seq. Primer: -28m13 rev2 ET from Amerham.

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

FEATURES  
 SOURCE  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="GDB:16039131"  
 /db\_xref="taxon:9606"  
 /clone\_id="IMAGE:795702"  
 /sex="male"  
 /lab\_host="DH1.0B"  
 /note="Scares testis\_NHT"  
 /polylinker: Site 1: Not I; Site 2: Eco RI; 1st strand cDNA  
 was prepared from mRNA obtained from Clontech  
 Laboratories, Inc., and primed with a Not I - oligo(dT)  
 primer [5].  
 Double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not I  
 and Eco RI sites of the modified PT733 vector. Library  
 went through one round of normalization to Cot5, and was  
 constructed by Bento Soares and M. Fatima Bonaldo.

ORIGIN  
 Query Match 14.6%; Score 146.4; DB 9; Length 390;  
 Best Local Similarity 98.8%; Pred. No. 1.e-25;  
 Matches 158; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 471 TTACGGATGTGCAATTGTAGACTTTGAGCACTTCATAATGCCTGTGAAGAGAAA 530  
 Db 244 TTTCAGGATGTGCAATTGTAAAGCACTTGAGCACTTCATAATGCCTGTGAAGAGAAA 185

Qy 531 TGTACCCAGATGTATCATTACCTTGTGCTGAGGCCGCTCTTTCAGGATTTCAGT 590  
 Db 184 TGTACCCAGATGTATCATTACCTTGTGCTGAGGCCGCTCTTTCAGGATTTCAGT 126

Qy 591 CACATCTCCCTGCTTGTGCAAGAACACATGACCAAGCTC 630  
 Db 125 CACATCTCCCTGCTTGTGCAAGAACACATGACCAAGCTC 86

RESULT 12  
 CC325294 LOCUS CC325294 533 bp mRNA linear GSS 14-MAY-2003  
 DEFINITION XM17 BayGenomics Gene Trap Library PGT0Lxf Mus musculus cDNA, mRNA  
 ACCESSION CC325294  
 VERSION CC325294.1 GI:30719352  
 KEYWORDS SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 533)  
 AUTHORS BayGenomics  
 TITLE http://baygenomics.ucsf.edu/  
 JOURNAL Unpublished (2003)

RESULT 11  
 AA459832 LOCUS AA459832 390 bp mRNA linear EST 09-JUN-1997  
 DEFINITION zx50g4\_r1 Scares testis\_NHT Homo sapiens cDNA clone IMAGE:795702  
 ACCESSION AA459832  
 VERSION AA459832.1 GI:2184739  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

COMMENT Contact: BayGenomics  
 Bay Area Functional Genomics Consortium (BayGenomics)  
 Email: info@baygenomics.ucsf.edu  
 Sequence tag generated by 5' RACE of total RNA from gene trap ES  
 Cell line. ES cell lines harboring insertion mutation of target  
 gene are available upon request from BayGenomics. Annotation  
 information available from  
[http://baygenomics.ucsf.edu/cgi-bin/BaySearch.py?OPTION=EXACT&TYPE=CELL\\_LINE&KEY=XM0179](http://baygenomics.ucsf.edu/cgi-bin/BaySearch.py?OPTION=EXACT&TYPE=CELL_LINE&KEY=XM0179)

FEATURES Class: Gene Trap;  
 Location/Qualifiers 1. .533  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="129 OLA"  
 /db\_xref="taxon:10090"  
 /sex="Male"  
 /cell\_type="Embryonic stem cell"  
 /clone\_lib="BayGenomics Gene Trap Library pGT0Lxf"  
 /note="Vector: pGT0Lxf"

ORIGIN Query Match 14.1%; Score 141.2; DB 28; Length 533;  
 Best Local Similarity 87.1%; Pred. No. 2.6e-24;  
 Matches 155; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 467 GAGATTACAGCAGATGCACTTGAAGCACATTGGGCCACTATAAATGCTGTGAGA 526  
 Db 103 GAGCACAGAACGATGCAATTGGCAAACACCTTGGGCCACTATAAATGCTGTGAGA 162

RESULT 14 CC325312 LOCUS CC325312 571 bp mRNA linear GSS 14-MAY-2003  
 DEFINITION XM085 BayGenomics Gene Trap Library pGT0Lxf Mus musculus cDNA, mRNA sequence.  
 ACCESSION CC325312  
 VERSION 1 GI:30719370  
 KEYWORDS GSS  
 SOURCE Mus musculus (house mouse)  
 ORGANISM  
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 51)  
 AUTHORS BayGenomics.  
 TITLE http://baygenomics.ucsf.edu/  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: BayGenomics Consortium (BayGenomics)  
 Bay Area Functional Genomics Consortium (BayGenomics)  
 Email: info@baygenomics.ucsf.edu  
 Sequence tag generated by 5' RACE of total RNA from gene trap ES  
 cell line. ES cell lines harboring insertion mutation of target  
 gene are available upon request from BayGenomics. Annotation  
[http://baygenomics.ucsf.edu/cgi-bin/BaySearch.py?OPTION=EXACT&TYPE=CELL\\_LINE&KEY=XM085](http://baygenomics.ucsf.edu/cgi-bin/BaySearch.py?OPTION=EXACT&TYPE=CELL_LINE&KEY=XM085).

FEATURES Source 1..571  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="129 OLA"  
 /db\_xref="taxon:10090"  
 /sex="Male"  
 /clone\_lib="BayGenomics Gene Trap Library pGT0Lxf"  
 /note="Vector: pGT0Lxf"

ORIGIN Query Match 14.1%; Score 141.2; DB 28; Length 571;  
 Best Local Similarity 87.1%; Pred. No. 2.6e-24;  
 Matches 155; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 467 GAGATTACAGCAGATGCACTTGAAGCACATTGGGCCACTATAAATGCTGTGAGA 526  
 Db 112 GAGCACAGAACGATGCAATTGGCAAACACCTTGGGCCACTATAAATGCTGTGAGA 171

RESULT 13 CC325314 LOCUS CC325314 560 bp mRNA linear GSS 14-MAY-2003  
 DEFINITION XM083 BayGenomics Gene Trap Library pGT0Lxf Mus musculus cDNA, mRNA sequence.  
 ACCESSION CC325314  
 VERSION 1 GI:30719372  
 KEYWORDS GSS  
 SOURCE Mus musculus (house mouse)  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 560)  
 AUTHORS BayGenomics.  
 TITLE http://baygenomics.ucsf.edu/  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: BayGenomics Consortium (BayGenomics)  
 Bay Area Functional Genomics Consortium (BayGenomics)  
 Email: info@baygenomics.ucsf.edu  
 Sequence tag generated by 5' RACE of total RNA from gene trap ES  
 cell line. ES cell lines harboring insertion mutation of target  
 gene are available upon request from BayGenomics. Annotation  
[http://baygenomics.ucsf.edu/cgi-bin/BaySearch.py?OPTION=EXACT&TYPE=CELL\\_LINE&KEY=XM083](http://baygenomics.ucsf.edu/cgi-bin/BaySearch.py?OPTION=EXACT&TYPE=CELL_LINE&KEY=XM083).

FEATURES Source 1..560  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="129 OLA"  
 /db\_xref="taxon:10090"  
 /sex="Male"  
 /clone\_lib="BayGenomics Gene Trap Library pGT0Lxf"  
 /note="Vector: pGT0Lxf"

ORIGIN Query Match 14.1%; Score 141.2; DB 28; Length 586;  
 Best Local Similarity 87.1%; Pred. No. 2.6e-24;  
 Matches 155; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 527 GAAATGTAACCGAGATGATCATTGTCGTTGAGCACATTGGGCCACTATAAATGCTGTGAGA 586  
 Db 172 GAGGTGACCCGAGATGACCATTCACATGTCGAGGGCTGCACTTTCAGACTT 231

RESULT 14 CC325314 LOCUS CC325314 587 bp mRNA linear GSS 14-MAY-2003  
 DEFINITION XM084 BayGenomics Gene Trap Library pGT0Lxf Mus musculus cDNA, mRNA sequence.  
 ACCESSION CC325314  
 VERSION 1 GI:30719370  
 KEYWORDS GSS  
 SOURCE Mus musculus (house mouse)  
 ORGANISM  
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 51)  
 AUTHORS BayGenomics.  
 TITLE http://baygenomics.ucsf.edu/  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: BayGenomics Consortium (BayGenomics)  
 Bay Area Functional Genomics Consortium (BayGenomics)  
 Email: info@baygenomics.ucsf.edu  
 Sequence tag generated by 5' RACE of total RNA from gene trap ES  
 cell line. ES cell lines harboring insertion mutation of target  
 gene are available upon request from BayGenomics. Annotation  
[http://baygenomics.ucsf.edu/cgi-bin/BaySearch.py?OPTION=EXACT&TYPE=CELL\\_LINE&KEY=XM084](http://baygenomics.ucsf.edu/cgi-bin/BaySearch.py?OPTION=EXACT&TYPE=CELL_LINE&KEY=XM084).

FEATURES Source 1..587  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="129 OLA"  
 /db\_xref="taxon:10090"  
 /sex="Male"  
 /clone\_lib="BayGenomics Gene Trap Library pGT0Lxf"  
 /note="Vector: pGT0Lxf"

Db	232	TAGTCACTTCTTCCTCTGTCAGAACACATCGACCAAGCTCTGAAAGATCAAAG	289		Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
RESULT	15			COMMENT	CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
AK054270	AK054270	Mus musculus 2 days pregnant adult female ovary cDNA, RIKEN full-length enriched library, clone: 333008K23 product: hypothetical HMG-I and HMG-Y DNA-binding domain (A+T-hook) PHD-finger/DHHC-type Zn-finger/RING finger containing protein, full insert sequence.			Please visit our web site for further details. URL: http://genome.gsc.riken.go.jp/
ACCESSION	AK054270	GI: 26144110		FEATURES	1. 2647 Location/Qualifiers
VERSION	HTC	CAP trapper.		SOURCE	Mus musculus (house mouse)
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE	2				
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.			CDS	
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
MEDLINE	10499374				
PUBMED	11042159				
REFERENCE	3				
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasakai, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Harada, M., Sumi, N., Ishii, Y., Nakamura, S., Hachisu, M., Nishime, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanuki, M., Yoneda, Y., Ishikawa, T., Ozawa, X., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.				
TITLE	RIKEN integrated sequence analysis (RISA) system -3.84-format sequencing pipeline with 384 multicapillary sequencer				
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)				
MEDLINE	11076861				
REFERENCE	4				
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the RIKEN Consortium.				
TITLE	Functional annotation of a full-length mouse cDNA collection				
JOURNAL	Nature 409, 685-690 (2001)				
AUTHORS	The RIKEN Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.				
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs				
JOURNAL	Nature 420, 563-573 (2002)				
REFERENCE	6 (bases 1 to 2647)				
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Hori, F., Furuno, M., Hanazono, K., Hiraoka, T., Hizukuri, T., Hayashida, K., Hayatsu, N., Hiramoto, K., Hara, A., Hashizume, W., Katoch, H., Kawai, J., Kohji, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Onsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shirasaki, T., Tagami, M., Tagawa, A., Takahashi, F., Tomaru, A., Toda, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.				
TITLE	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Japan				
JOURNAL	Direct Submission				

Search completed: July 9, 2004, 10:07:38  
Job time : 4200 secs

This Page Blank (uspto)

GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2004, 16:57:41 ; Search time 55 Seconds  
 (without alignments)

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:  
 1: geneseqD1980s:  
 2: geneseqD1990s:  
 3: geneseqD000s:  
 4: geneseqD2001s:  
 5: geneseqD2002s:  
 6: geneseqD2003as:  
 7: geneseqD2003bs:  
 8: geneseqD2004s:  
 8: geneseqD2004as:  
 Total number of hits satisfying chosen parameters: 1586107

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	213	100.0	43	2	AAW02152	BAGE tumo	Aaw02152 BAGE tumo
2	213	100.0	43	5	AAU84811	Human BAG	Aau84811 Human BAG
3	213	100.0	43	5	ABB78347	Amino aci	Abb78347 Amino aci
4	213	100.0	43	7	ADD25520	Birding d	Add25520 Birding d
5	156	73.2	30	5	AAU85010	Human BAG	Aau85010 Human BAG
6	156	73.2	30	5	AAU85130	Human mel	Aau85130 Human mel
7	126	59.2	30	5	AAU85509	Human BAG	Aau85509 Human BAG
8	98	46.0	22	2	AAU67808	BACE tumo	Aau67808 BACE tumo
9	98	46.0	22	2	AAU10634	Peptide a	Aau10634 Peptide a
10	98	46.0	22	5	ABG80319	MHC Class	Abg80319 MHC Class
11	87	40.8	17	5	AAU85011	Human BAG	Aau85011 Human BAG
12	70	32.9	16	2	AAU67809	BAGE tumo	Aau67809 BAGE tumo
13	70	32.9	16	2	AAV10635	Peptide a	Aav10635 Peptide a
14	70	32.9	16	5	ABG80316	MHC Class	Abg80316 MHC Class
15	63	29.6	384	4	AAU8007	Human sph	Aau8007 Human sph
16	63	29.6	384	6	ABP71054	Human sph	Abp71054 Human sph
17	62	29.1	403	13	AAW84613	Secreted	Aaw84613 Secreted
18	62	29.1	293	4	ABBS0310	Human sec	Abbs0310 Human sec
19	62	29.1	293	6	ABO4637	Novel hum	Abo4637 Novel hum
20	62	29.1	293	7	ABO26117	Human pro	Abo26117 Human pro
21	62	29.1	305	4	AAU1966	Human pol	Aau1966 Human pol
22	62	29.1	333	5	ABBD0209	Human pol	Abbd0209 Human pol
23	62	29.1	368	6	ABR2291	Human sph	Abr2291 Human sph
24	62	29.1	384	3	ABJ18659	Human r	Abj18659 Human r
25	62	29.1	384	3	AYA96057	Human sph	Aya96057 Human sph

RESULT 1  
 AAW02152  
 ID AAW02152 standard; Protein; 43 AA.  
 XX  
 AC  
 XX  
 DT 04-DEC-1996 (First entry)  
 XX  
 DE BAGE tumour rejection antigen precursor.  
 XX  
 BAGE; tumour rejection antigen precursor; TRAP; MHC;  
 KW major histocompatibility complex; HLA-Cw\*1601; melanoma; metastasis;  
 KW diagnosis; therapy; vaccine.  
 XX  
 Homo sapiens.  
 XX  
 CS  
 XX  
 PH  
 Key Peptide  
 FT 2...10  
 FT /label= Tumour\_rejection\_antigen  
 FT  
 XX  
 EN  
 WO9625511-A1.  
 XX  
 PD 22-AUG-1996.  
 XX  
 PP 07-FEB-1996;  
 96WO-US001608.  
 XX  
 PR 16-FEB-1995;  
 95US-00389360.  
 XX  
 PA (LUDWIG INST CANCER RES.  
 XX  
 PI Boel P, Wildmann C, Boonfalleur T, Van Der Bruggen P, Coulie P;  
 PI Renaud J;  
 XX  
 DR WPI: 1996-393411/39.  
 DR N-PSDB; AAT36332.  
 XX  
 PT Tumour rejection antigen precursor (TRAP) and gene - useful to develop prods. for diagnosis and treatment of disorders characterised by TRAP, partic. melanoma.  
 PT  
 XX  
 PS Example 4; Page 11; 44DP; English.  
 XX  
 CC The BAGE tumour rejection antigen precursor (TRAP) (AAW02152) is processed to a tumour rejection antigen (TRA) (AAW02153) presented by the MHC molecule HLA-Cw1601. The amino acid sequence of the BAGE TRAP was deduced from a cDNA clone (AAT36382) derived from melanoma cell line MZ2-MEL 43. BAGE expression was not observed in healthy adult or foetal tissue, but was detected in 22% of melanoma lines examined, being partic.

CC prevalent in metastatic lesions. The BAGE antigen can be expressed in transformed or transfected host cells. It is useful for identifying those cells which present HLA-Cw1601 and may be incorporated into vaccines useful in treating disorders characterised by expression of the TRAP molecule

XX Sequence 43 AA;

Query Match	100.0%	Score 213;	DB 2;	Length 43;
Best Local Similarity	100.0%;	Pred. No. 1.	8e-23;	
Matches	43;	Conservative	0;	Gaps 0;
Qy	1 MAARAVFLAQLAQQLMKKEESPVVSWRLPEDGTALCPIF 43			
Db	1 MAARAVFLAQLAQQLMKKEESPVVSWRLPEDGTALCPIF 43			

RESULT 2  
AAU84811

ID AAU84811 standard; protein; 43 AA.

XX

AC AAU84811;

XX

DT 08-MAY-2002 (first entry)

XX

DE Human BAGE consensus sequence.

XX

KW Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;

KW viral infection; human immunodeficiency virus; melanoma;

KW bacterial infection; Salmonella; Legionella; parasitic infection;

KW Trypanosoma; Toxoplasma; Giardia.

XX

OS Homo sapiens.

XX

PN WO200190197-A1.

XX

PD 29-NOV-2001.

XX

PP 25-MAY-2001; 2001WO-AU0006622.

XX

PR 26-MAY-2000; 2000AU\_00007761.

XX

(AUSU ) UNIV AUSTRALIAN NAT.

XX

PI Thomson SA, Ramshaw IA;

XX

DR WPI; 2002-147575/19.

XX

PT New synthetic polypeptides having several different segments of at least one parent polypeptide linked together differently compared to the linkage in the parent polypeptide, for inducing immune response against a pathogen or cancer.

XX

PT Example 3; Fig 27; 364pp; English.

XX

PS The invention relates to a new synthetic polypeptide (I) comprising several different segments of at least one parent polypeptide linked together in a different relationship relative to their linkage in the parent polypeptide to impede, abrogate or otherwise alter at least one function associated with the parent polypeptide and for inducing an immune response against a pathogen or cancer. Also included are a synthetic polynucleotide encoding and a computer system for designing the synthetic polypeptides. The synthetic polypeptides and polynucleotides are referred to as a Savine. The synthetic polypeptide is useful for modulating immune responses preferably directed against a pathogen or a cancer, (e.g., cancers of the lung, breast, cervix, colon, head and neck, pancreas, prostate, stomach, bladder, kidney, bone liver, oesophagus, brain, testicle, uterus), as parentential agents.

CC Compositions comprising the polypeptide may be used in the treatment or prophylaxis against viral (such as infections caused by HIV (human immunodeficiency virus), hepatitis, influenza, Japanese encephalitis virus, Epstein-Barr virus and respiratory syncytial virus), bacterial (e.g., infections caused by Neisseria, Meningococcal, Haemophilus,

CC Salmonella, Streptococcal, Legionella and Mycobacterium or parasitic (e.g., infections caused by Plasmodium, Schistosoma, Leishmania, Trypanosoma, Toxoplasma and Giardia) infections. The present sequence is a consensus sequence for a parent protein used to design a Savine of the invention

CC Sequence 43 AA;

CC SQ

Query Match	100.0%;	Score 213;	DB 5;	Length 43;
Best Local Similarity	100.0%;	Pred. No. 1.	8e-23;	
Matches	43;	Conservative	0;	Gaps 0;
Qy	1 MAARAVFLAQLAQQLMKKEESPVVSWRLPEDGTALCPIF 43			
Db	1 MAARAVFLAQLAQQLMKKEESPVVSWRLPEDGTALCPIF 43			

RESULT 3

ABB73347

ID ABB73347 standard; protein; 43 AA.

XX

AC ABB73347;

XX

DT 16-DEC-2002 (first entry)

XX

DE Amino acid sequence of human BAGE.

XX

KW Human; IFIT-2; chronic myelogenous leukemia; IAGE-1; BAGE; DDB1; ETS2; PIASy; PIASx-alpha; PIASx-beta; DAPK3.

XX

OS Homo sapiens.

XX

PN WO200270747-A1.

XX

PD 12-SSP-2002.

XX

PF 01-MAR-2002; 2002WO-JP001901.

XX

PR 01-MAR-2001; 2001JP-000056438.

XX

PA (FUJII ) FUJISAWA PHARM CO LTD.

XX

PI Mano H;

XX

WPI; 2002-682911/73.

XX

DR N-PSB; ABV72225.

XX

PT Measuring the expression profile of Genes in a cell or tissue sample for diagnosis of chronic myelogenous leukemia and identification of agents for its treatment.

XX

PS Example 3; Page 39-40; 97pp; Japanese.

XX

CC The present sequence represents human BAGE. The expression level of the gene is used in the method of the invention. The specification describes a method of examining chronic myelogenous leukemia. The method comprises measuring the expression level of a gene selected from IFIT-2, LAGE-1, CC BAGE, DDB1, ETS2, PIASy, PIASx-alpha, PIASx-beta and DAPK3, or CC determining the expression profile of a group of genes including one or CC more of these genes, in a cell or tissue sample from a chronic CC myelogenous leukemia patient. The method is used for the diagnosis, CC treatment and prevention of chronic myelogenous leukemia

XX

SQ Sequence 43 AA;

CC Query Match

100.0%;

Score 213;

DB 5;

Length 43;

Best Local Similarity

100.0%;

Pred. No. 1.

8e-23;

Mismatches

0;

Indels

0;

Gaps

0;

PS Example 3; Page 39-40;

CC Query Match

100.0%;

Score 213;

DB 5;

Length 43;

Best Local Similarity

100.0%;

Pred. No. 1.

8e-23;

Mismatches

0;

Indels

0;

Gaps

0;

PS Example 3; Page 39-40;

CC Query Match

100.0%;

Score 213;

DB 5;

Length 43;

Best Local Similarity

100.0%;

Pred. No. 1.

8e-23;

Mismatches

0;

Indels

0;

Gaps

0;

PS Example 3; Page 39-40;

CC Query Match

100.0%;

Score 213;

DB 5;

Length 43;

Best Local Similarity

100.0%;

Pred. No. 1.

8e-23;

Mismatches

0;

Indels

0;

Gaps

0;

**RESULT 4**  
**ADD2520 standard; protein; 43 AA.**  
**ID ADD2520;**  
**XX**  
**AC**  
**XX**  
**DT 15-JAN-2004 (first entry)**  
**DE Binding domain-immunoglobulin fusion protein-associated protein #37.**  
**XX**  
**KW Binding domain; immunoglobulin; fusion Protein; cytosatic;**  
**KW antiarthritic; immunosuppressive; antidiabetic; antithyroid;**  
**KW neuroprotective; hinge region; immunoglobulin heavy chain;**  
**KW CH2 constant region; CH3 constant region; IgG1;**  
**KW antibody dependent cell-mediated cytotoxicity; ADCC; complement fixation;**  
**KW malignant condition; B-cell disorder; melanoma; carcinoma; sarcoma;**  
**KW rheumatoid arthritis; myasthenia gravis; Grave's disease;**  
**KW type I diabetes mellitus; multiple sclerosis; autoimmune disease.**  
**XX**  
**OS Unidentified.**  
**PN US2003118592-A1.**  
**XX**  
**PD 26-JUN-2003.**  
**XX**  
**PF 25-JUL-2002; 2002US-00207655.**  
**XX**  
**PR 17-JAN-2001; 2001US-0367358P.**  
**PR 17-JAN-2002; 2002US-00053350.**  
**PR 03-JUN-2002; 2002US-0385691P.**  
**XX**  
**PA (GENE-) GENECRAFT INC.**

**PI Ledbetter JA, Hayden-Ledbetter MS, Thompson PA;**  
**XX**  
**DR 2003-801317/75.**  
**XX**  
**PT New binding domain-immunoglobulin fusion protein, useful for treating a subject having or suspected of having a malignant condition or a B-cell disorder, e.g. melanoma, Grave's disease or autoimmune disease.**  
**XX**  
**PS Disclosure: SEQ ID NO 81; 157pp; English.**  
**XX**

The invention relates to a binding domain-immunoglobulin fusion protein comprising a binding domain polypeptide that is fused to an immunoglobulin hinge region polypeptide, an immunoglobulin heavy chain CH2 constant region polypeptide that is fused to the hinge region polypeptide, and an immunoglobulin heavy chain CH3 constant region polypeptide that is fused to the CH2 constant region polypeptide. The hinge region polypeptide comprises: a wild-type human IgG1 hinge region polypeptide, a mutated human IgG1 immunoglobulin hinge region polypeptide, derived from (a) having 3 or more cysteine residues; where the mutated human IgG1 immunoglobulin hinge region polypeptide contains no cysteine residues, where the first cysteine is not mutated; a mutated human IgG1 immunoglobulin hinge region polypeptide, derived from (a) having 3 or more cysteine residues, where the mutated human IgG1 immunoglobulin hinge region polypeptide contains no more than one cysteine residue; and a mutated human IgG1 immunoglobulin hinge region polypeptide comprising antibody dependent cell-mediated cytotoxicity (ADCC) and complement fixation. The binding domain polypeptide is capable of specifically binding to an antigen. Also included are an isolated polynucleotide encoding the binding domain-immunoglobulin fusion protein, a recombinant expression construct comprising the polynucleotide operably linked to a promoter, a host cell transformed or transfected with a recombinant expression construct, producing the binding domain-immunoglobulin fusion protein, a pharmaceutical composition comprising the binding domain-immunoglobulin fusion protein or polynucleotide and a carrier, and treating a subject having or suspected of having a malignant condition or a B-cell disorder.

**CC The binding domain-immunoglobulin fusion protein is useful for treating a subject having or suspected of having a malignant condition or a B-cell disorder, e.g. melanoma, carcinoma or sarcoma, rheumatoid arthritis, myasthenia gravis, Grave's disease, Type I diabetes mellitus, multiple sclerosis or autoimmune disease. The present sequence is a binding domain -immunoglobulin fusion protein-associated protein sequence. Note: The sequence data for this patent formed part of the printed specification and is also available in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html?DocID=20030118592. The authors have not identified the sequences in the printed specification by their SEQ ID number therefore none of the sequences can be explicitly identified.**  
**XX**  
**SQ Sequence 43 AA;**  
**Query Match 100.0%; Score 213; DB 7; Length 43;**  
**Best Local Similarity 100.0%; Pred. No. 1.8e-23;**  
**Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;**  
**Qy 1 MAARAVFLAQLAQQLQARLMKEESPVSVWLEPEDTALCFIF 43**  
**Db 1 MAARAVFLAQLAQQLQARLMKEESPVSVWLEPEDTALCFIF 43**  
**XX**  
**RESULT 5**  
**AAU85010**  
**ID AAU85010 standard; peptide; 30 AA.**  
**XX**  
**XX**  
**AC AAU85010;**  
**XX**  
**DT 08-MAY-2002 (first entry)**  
**DE Human BAGE segment 2.**  
**XX**  
**Saving; vaccine; cancer; viral infection; HIV; hepatitis C virus;**  
**KW viral infection; human immunodeficiency virus; melanoma;**  
**KW bacterial infection; Salmonella; Legionella; Toxoplasma; parasitic infection;**  
**XX**  
**OS Homo sapiens.**  
**XX**  
**FN WO200190197-A1.**  
**XX**  
**PP 25-MAY-2001; 2001WO-AU000622.**  
**XX**  
**PR 26-MAY-2000; 2000AU-0007761.**  
**XX**  
**PA (AUSU ) UNIV AUSTRALIAN NAT.**  
**XX**  
**PI Thomson SA, Ramshaw IA;**  
**XX**  
**WPI; 2002-14755/19.**  
**XX**  
**DR N-PSDB; ABK368-0.**  
**XX**  
**PT New synthetic polypeptides having several different segments of at least one parent polypeptide linked together differently compared to the linkage in the parent polypeptide, for inducing immune response against a pathogen or cancer.**  
**XX**  
**PS Example 3; Fig 27; 364pp; English.**  
**XX**  
**PT The invention relates to a new synthetic polypeptide (I) comprising several different segments of at least one parent polypeptide linked together in a different linkage relative to their linkage in the parent polypeptide to impede, abrogate or otherwise alter at least one function associated with the parent polypeptide and for inducing an immune response against a pathogen or cancer. Also included are a synthetic polynucleotide encoding a computer system for designing the synthetic polypeptides. The synthetic polypeptides and polynucleotides are referred to as a Savine. The synthetic polypeptide is useful for modulating immune responses preferentially directed against a pathogen or a cancer, (e.g., cancers of the lung, breast, ovary, cervix, colon, head**  
**CC**

CC and neck, pancreas, prostate, stomach, bladder, kidney, bone liver, oesophagus, brain, testicle, uterus), as potentiating agents.

CC Compositions comprising the polypeptide may be used in the treatment or prophylaxis against viral (such as infections caused by HRV (human immunodeficiency virus), hepatitis, influenza, Japanese encephalitis (e.g., infections caused by *Neisseria*, *Meningococcal*, *Haemophilus*, *Salmonella*, *Streptococcal*, *Legionella* and *Mycobacterium* or parasitic (e.g., infections caused by *Plasmodium*, *Schistosoma*, *Leishmania*, *Trypanosoma*, *Toxoplasma* and *Giardia*) infections. The present sequence is a peptide derived from a parent protein used to construct a Savine of the invention

XX Sequence 30 AA;

Query Match 73.2%; Score 156; DB 5; Length 30;  
Best Local Similarity 100.0%; Pred. No. 2.5e-15; Indels 0; Gaps 0;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 LLQARLMKESSESPVYSWRLSPEDGTALCFIF 43  
Db 1 LLQARLMKESSESPVYSWRLSPEDGTALCFIF 30

## RESULT 6

AAU85130

ID AAU85130 standard; protein; 3541 AA.

XX

AC AAU85130;

XX

DT 08-MAY-2002 (first entry)

XX Human melanoma specific savine.

KW Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus; viral infection; human immunodeficiency virus; melanoma; bacterial infection; *Salmonella*; *Legionella*; parasitic infection; *Trypanosoma*; *Toxoplasma*; *Giardia*.

OS Homo sapiens.

OS Synthetic.

PN WO200190197-A1.

XX PD 29-NOV-2001.

XX PF 25-MAY-2001; 2001WO-AU000622.

XX PR 26-MAY-2000; 2000AU-00007761.

XX PA (AUSU ) UNIV AUSTRALIAN NAT.

PA Thomson SA, Ramshaw IA;

XX PI Thomson SA, Ramshaw IA;

DR WPI; 2002-147575/19.

DR N-PSDB; ABK36950.

XX New synthetic polypeptides having several different segments of at least one parent polypeptide linked together differently compared to the linkage in the parent polypeptide, for inducing immune response against a pathogen or cancer.

XX PS Example 3; Fig 27; 364DP; English.

XX The invention relates to a new synthetic polypeptide (I) comprising several different segments of at least one parent polypeptide linked together in a different relationship relative to their linkage in the parent polypeptide to impede, abrogate or otherwise alter at least one function associated with the parent polypeptide and for inducing an immune response against a pathogen or cancer. Also included are a synthetic polynucleotide encoding and a computer system for designing the synthetic polypeptides. The synthetic polypeptides and polynucleotides are referred to as a Savine. The synthetic polypeptide is useful for

CC modulating immune responses preferably directed against a pathogen or a cancer, (e.g., cancers of the lung, breast, ovary, cervix, colon, head and neck, pancreas, prostate, stomach, bladder, kidney, bone liver, oesophagus, brain, testicle, uterus) as potentiating agents.

CC Compositions comprising the polypeptide may be used in the treatment or prophylaxis against viral (such as infections caused by HIV (human immunodeficiency virus), hepatitis, influenza, Japanese encephalitis (e.g., infections caused by *Neisseria*, *Meningococcal*, *Haemophilus*, *Salmonella*, *Streptococcal*, *Legionella*, *Mycobacterium* or parasitic (e.g., infections caused by *Neisseria*, *Meningococcal*, *Hemophilus*, *Salmonella*, *Streptococcal*, *Legionella* and *Mycobacterium* or parasitic (e.g., infections caused by *Plasmodium*, *Schistosoma*, *Leishmania*, *Trypanosoma*, *Toxoplasma* and *Giardia*) infections. The present sequence is a peptide derived from a parent protein used to construct a Savine of the invention

XX a Savine protein of the invention

XX Sequence 30 AA;

Query Match 73.2%; Score 156; DB 5; Length 3541;  
Best Local Similarity 100.0%; Pred. No. 6.1e-13; Indels 0; Mismatches 0; Gaps 0;

Qy 14 LLQARLMKESSESPVYSWRLSPEDGTALCFIF 43  
Db 2777 LLQARLMKESSESPVYSWRLSPEDGTALCFIF 2806

## RESULT 7

AAU85009

ID AAU85009 standard; peptide; 30 AA.

XX AAU85009;

AC AAU85009;

XX DT 08-MAY-2002 (first entry)

XX Human BAGE segment 1.

KW Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus; viral infection; human immunodeficiency virus; melanoma; bacterial infection; *Salmonella*; *Legionella*; parasitic infection; *Trypanosoma*; *Toxoplasma*; *Giardia*.

OS Homo sapiens.

OS Homo sapiens.

PN WO200190197-A1.

XX PD 29-NOV-2001.

XX PF 25-MAY-2001; 2001WO-AU000622.

XX PR 26-MAY-2000; 2000AU-00007761.

XX PA (AUSU ) UNIV AUSTRALIAN NAT.

PA Thomson SA, Ramshaw IA;

XX PI Thomson SA, Ramshaw IA;

DR WPI; 2002-147575/19.

DR N-PSDB; ABK36950.

XX New synthetic polypeptides having several different segments of at least one parent polypeptide linked together differently compared to the linkage in the parent polypeptide, for inducing immune response against a pathogen or cancer.

XX PS Example 3; Fig 27; 364DP; English.

XX The invention relates to a new synthetic polypeptide (I) comprising several different segments of at least one parent polypeptide linked together in a different relationship relative to their linkage in the parent polypeptide to impede, abrogate or otherwise alter at least one function associated with the parent polypeptide and for inducing an immune response against a pathogen or cancer. Also included are a synthetic polynucleotide encoding and a computer system for designing the synthetic polypeptides. The synthetic polypeptides and polynucleotides are referred to as a Savine. The synthetic polypeptide is useful for

CC modulating immune responses preferably directed against a pathogen or a cancer, (e.g., cancers of the lung, breast, ovary, cervix, colon, head and neck, pancreas, prostate, stomach, bladder, kidney, bone liver, oesophagus, brain, testicle, uterus) as potentiating agents.

CC Compositions comprising the polypeptide may be used in the treatment or prophylaxis against viral (such as infections caused by HIV (human immunodeficiency virus), hepatitis, influenza, Japanese encephalitis (e.g., infections caused by *Neisseria*, *Meningococcal*, *Haemophilus*, *Salmonella*, *Streptococcal*, *Legionella*, *Mycobacterium* or parasitic (e.g., infections caused by *Neisseria*, *Meningococcal*, *Hemophilus*, *Salmonella*, *Streptococcal*, *Legionella* and *Mycobacterium* or parasitic (e.g., infections caused by *Plasmodium*, *Schistosoma*, *Leishmania*, *Trypanosoma*, *Toxoplasma* and *Giardia*) infections. The present sequence is a Savine protein of the invention

XX a Savine protein of the invention

XX Sequence 3541 AA;

XX Query Match 73.2%; Score 156; DB 5; Length 3541;

Best Local Similarity 100.0%; Pred. No. 6.1e-13; Indels 0; Mismatches 0; Gaps 0;

Qy 14 LLQARLMKESSESPVYSWRLSPEDGTALCFIF 43  
Db 2777 LLQARLMKESSESPVYSWRLSPEDGTALCFIF 2806

## RESULT 7

AAU85009

ID AAU85009 standard; peptide; 30 AA.

XX AAU85009;

AC AAU85009;

XX DT 08-MAY-2002 (first entry)

XX Human BAGE segment 1.

KW Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus; viral infection; human immunodeficiency virus; melanoma; bacterial infection; *Salmonella*; *Legionella*; parasitic infection; *Trypanosoma*; *Toxoplasma*; *Giardia*.

OS Homo sapiens.

OS Homo sapiens.

PN WO200190197-A1.

XX PD 29-NOV-2001.

XX PF 25-MAY-2001; 2001WO-AU000622.

XX PR 26-MAY-2000; 2000AU-00007761.

XX PA (AUSU ) UNIV AUSTRALIAN NAT.

PA Thomson SA, Ramshaw IA;

XX PI Thomson SA, Ramshaw IA;

DR WPI; 2002-147575/19.

DR N-PSDB; ABK36950.

XX New synthetic polypeptides having several different segments of at least one parent polypeptide linked together differently compared to the linkage in the parent polypeptide, for inducing immune response against a pathogen or cancer.

XX PS Example 3; Fig 27; 364DP; English.

XX The invention relates to a new synthetic polypeptide (I) comprising several different segments of at least one parent polypeptide linked together in a different relationship relative to their linkage in the parent polypeptide to impede, abrogate or otherwise alter at least one function associated with the parent polypeptide and for inducing an immune response against a pathogen or cancer. Also included are a synthetic polynucleotide encoding and a computer system for designing the synthetic polypeptides. The synthetic polypeptides and polynucleotides are referred to as a Savine. The synthetic polypeptide is useful for

CC modulating immune responses preferably directed against a pathogen or a cancer, (e.g., cancers of the lung, breast, ovary, cervix, colon, head and neck, pancreas, prostate, stomach, bladder, kidney, bone liver, oesophagus, brain, testicle, uterus) as potentiating agents.

CC Compositions comprising the polypeptide may be used in the treatment or prophylaxis against viral (such as infections caused by HIV (human immunodeficiency virus), hepatitis, influenza, Japanese encephalitis (e.g., infections caused by *Neisseria*, *Meningococcal*, *Haemophilus*, *Salmonella*, *Streptococcal*, *Legionella*, *Mycobacterium* or parasitic (e.g., infections caused by *Neisseria*, *Meningococcal*, *Hemophilus*, *Salmonella*, *Streptococcal*, *Legionella* and *Mycobacterium* or parasitic (e.g., infections caused by *Plasmodium*, *Schistosoma*, *Leishmania*, *Trypanosoma*, *Toxoplasma* and *Giardia*) infections. The present sequence is a Savine protein of the invention

XX a Savine protein of the invention

XX Sequence 3541 AA;

XX Query Match 73.2%; Score 156; DB 5; Length 3541;

Best Local Similarity 100.0%; Pred. No. 6.1e-13; Indels 0; Mismatches 0; Gaps 0;

Qy 14 LLQARLMKESSESPVYSWRLSPEDGTALCFIF 43  
Db 2777 LLQARLMKESSESPVYSWRLSPEDGTALCFIF 2806

CC modulating immune responses preferably directed against a pathogen or a cancer. (e.g. cancers of the lung, breast, ovary, cervix, colon, head and neck, pancreas, prostate, bladder, kidney, bone liver, oesophagus, brain, testicle, uterus), as potentiating agents.

CC Compositions comprising the polypeptide may be used in the treatment or prophylaxis against viral (such as infections caused by HIV (human immunodeficiency virus), hepatitis, influenza, Japanese encephalitis virus, Epstein-Barr virus and respiratory syncytial virus), bacterial (e.g., infections caused by Neisseria, Meningococcal, Haemophilus, Salmonella, Streptococcal, Legionella and Mycobacterium or parasitic (e.g., Trypanosoma, Toxoplasma and Giardia) infections. The present sequence is a peptide derived from a parent protein used to construct a savine of the invention

XX Sequence 30 AA;

Query Match 59.2%; Score 126; DB 5; Length 30;  
Best Local Similarity 100.0%; Pred. No. 6.1e-11;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAARAVFLAISQQLQARLMKE 28  
Db 3 MAARAVFLAISQQLQARLMKEESPVVS 30

RESULT 8

AAR67808 standard; peptide; 22 AA.  
XX

AC AAR67808;

XX DT 25-MAR-2003 (revised)

DT 22-AUG-1995 (first entry)

XX DE BAGE tumor rejection antigen peptide.

XX KW BAGE: tumor rejection antigen precursor; diagnosis; HLA;

KW human leukocyte antigen MHC; major histocompatibility complex; TRAP;

KW cancer; melanoma.

XX OS Synthetic.

XX PN WO9500159-A1.

XX PD 05-JAN-1995.

XX PF 10-JUN-1994; 94WO-US006534.

XX PR 17-JUN-1993; 93US-00079110.

XX 15-FEB-1994; 94US-00196630.

XX PA (LUDWIG INST CANCER RES.

XX PI Van Der Bruggen P, Boon-Falleur T, Coulie P, Renaud J;

XX DR WPI; 1995-051741/07.

XX Nucleic acid coding for a tumour rejection antigen precursor - used to develop prods. for the diagnosis and therapy of cancers, partic. melanomas.

XX Claim 21; Page 19; 33pp; English.

XX This sequence encodes the tumor rejection antigen peptide BAGE. The peptide may be used in the diagnosis and therapy of cancers, e.g. melanomas. (Updated on 25-MAR-2003 to correct BN field.)

XX SQ Sequence 22 AA;

XX Query Match 46.0%; Score 98; DB 2; Length 22;  
CC Best Local Similarity 100.0%; Pred. No. 5.2e-07;  
CC Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX SQ 1 MAARAVFLAISQQLQARLMKE 22

Db 1 MAARAVFLAISQQLQARLMKE 22

XX Query Match 46.0%; Score 98; DB 2; Length 22;  
CC Best Local Similarity 100.0%; Pred. No. 5.2e-07;  
CC Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 10

CC 1 MAARAVFLAISQQLQARLMKE 22  
CC 1 MAARAVFLAISQQLQARLMKE 22

DB RESULT 9  
AAV10634 standard; peptide; 22 AA.

XX ID AAV10634;

AC AAV10634;

DT 12-MAY-1999 (first entry)

XX DB Peptide antigen SEQ ID NO: 564.

XX KW Cytotoxic T-lymphocyte response; CTL; antigen; lymphatic system;

KW immunisation; tumour; infectious disease; immunotherapy; cancer; malignant melanoma; viral disease; hepatitis; AIDS.

XX OS Synthetic.

OS Homo sapiens.

XX PN WO9902183-A2.

XX PD 21-JAN-1999.

XX PF 10-JUL-1998; 98WO-US014289.

XX PR 10-JUL-1997; 97CA-02209815.

XX PR 10-DEC-1997; 97US-0098820.

XX PA (CTL)- CTL IMMUNOTHERAPIES CORP.

XX P1 Kuendig TM, Simard JJL;

XX DR WPI; 1999-120514/10.

XX Inducing a cytotoxic T lymphocyte response - by maintaining a level of PT antigen in the lymphatic system of a mammal so as to provide a sustained PT response, used to treat, e.g. AIDS.

XX Disclosure; Page 52; 1999pp; English.

XX PS DR WPI; 1999-120514/10.

CC The present invention describes a method of inducing and/or sustaining an immunological cytotoxic T lymphocyte (CTL) response in a mammal. The CC method comprises: (a) delivering an antigen to the mammal at a level to induce an immunological CTL response in the mammal; and (b) maintaining the CC level of the antigen in the mammal's lymphatic system to maintain the CC immunologic CTL response. The method can be used for the delivery of e.g. CC a differentiation antigen, a tumour-specific multilneage antigen, an CC embryonic antigen, an oncogene antigen, a mutated tumour-suppressor gene CC antigen, or a viral antigen. They can be used for the treatment of CC disease such as cancer, e.g. malignant melanoma or infectious disease, CC e.g. viral disease such as hepatitis or AIDS. Sustained antigen delivery CC to the lymphatic system provides for potent CTL stimulation that takes CC place in the milieu of the lymphoid organ, and its subsains stimulation CC that is necessary to keep CTL active, cytotoxic and recirculating through CC the body. AAV10071 to AAV10639 represent examples of peptide antigens CC given in the present invention

XX Sequence 22 AA;

Query Match 46.0%; Score 98; DB 2; Length 22;  
Best Local Similarity 100.0%; Pred. No. 5.2e-07;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAARAVFLAISQQLQARLMKE 22

Db 1 MAARAVFLAISQQLQARLMKE 22

XX Query Match 46.0%; Score 98; DB 2; Length 22;  
CC Best Local Similarity 100.0%; Pred. No. 5.2e-07;  
CC Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ABG80319 standard; peptide; 22 AA.  
 XX DB Human BAGE segment 3.  
 ID XX  
 ABG80319; XX  
 AC XX Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;  
 XX KW viral infection; human immunodeficienty virus; melanoma;  
 DT XX bacterial infection; Salmonella; Legionella; parasitic infection;  
 29-AUG-2003 (revised) KW Trypanosoma; Toxoplasma; Giardia.  
 DT 15-NOV-2002 (first entry) XX  
 XX  
**MHC class I molecule, viral epitope #567.**  
 KW Major histocompatibility complex; MHC class I molecule; virus;  
 KW epitope; cytotoxic T lymphocyte response; CTL response; lymphatic system;  
 KW antigen; immunogenic; malignant tumour; carcinoma; melanoma; leukaemia;  
 KW lymphoma; infectious disease; hepatitis; malaria; measles; tuberculosis;  
 KW acquired immune deficiency syndrome; AIDS.  
 OS Viruses.  
 XX Homo sapiens.  
 PN WO200190197-A1.  
 PA XX  
 PD 29-NOV-2001.  
 XX  
**MHC class I molecule, viral epitope #567.**  
 KW Major histocompatibility complex; MHC class I molecule; virus;  
 KW epitope; cytotoxic T lymphocyte response; CTL response; lymphatic system;  
 KW antigen; immunogenic; malignant tumour; carcinoma; melanoma; leukaemia;  
 KW lymphoma; infectious disease; hepatitis; malaria; measles; tuberculosis;  
 KW acquired immune deficiency syndrome; AIDS.  
 OS Viruses.  
 XX PA (AUSU ) UNIV AUSTRALIAN NAT.  
 PN PA  
 PD PI Thomson SA, Ramshaw IA;  
 XX PT  
 XX DR WPI: 2002-147575/19.  
 N-PSSDB; ABK36831.  
 DR XX  
**New synthetic polypeptides having several different segments of at least one parent polypeptide linked together differently compared to the parent polypeptide, for inducing immune response against a pathogen or cancer.**  
 PR PT New synthetic polypeptides having several different segments of at least one parent polypeptide linked together differently compared to the parent polypeptide, for inducing immune response against a pathogen or cancer.  
 PA PT  
 PI PT  
 XX PS Example 3; Fig 27; 364pp; English.  
 XX  
**Inducing or sustaining immunological cytotoxic T lymphocyte response in a mammal, useful for treating a mammal with malignant tumor or infectious disease, by directly administering an antigen to the lymphatic system of the mammal.**  
 XX Disclosure; Page 46; 73pp; English.  
 XX The invention relates to a method of inducing and/or sustaining an immunological cytotoxic T lymphocyte (CTL) response in a mammal comprising administering directly to the lymphatic system of the mammal:  
 CC (a) an antigen in the form of a polypeptide; (b) a vector comprising a nucleic acid encoding the antigen; or (c) a non peptide antigen. The method is useful for inducing and/or sustaining CTL response in a mammal. This is particularly useful for treating a mammal having a malignant tumour (e.g. carcinoma, melanoma, leukaemia or lymphoma) or infectious disease (e.g. hepatitis, acquired immune deficiency syndrome (AIDS), malaria, measles or tuberculosis), or in an animal having a predisposition to these diseases. The mammal may be dogs, cats, mice, cattle, sheep, pigs, goats, rabbits, or preferably humans. ABG19751-ABG80319 represent viral epitopes on major histocompatibility complex (MHC) class I molecules, used in the method of the invention. (Updated on 29-AUG-2003 to standardise OS field)  
 XX SQ Sequence 22 AA;  
 CC Query Match 46.0%; Score 98; DB 5; Length 22;  
 CC Best Local Similarity 100.0%; Pred. No. 5.2e-07; Mismatches 0; Indels 0; Gaps 0;  
 CC Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC  
 Qy 1 MAARAVFLNSAQQLQARLMKE 22  
 Db 1 MAARAVFLNSAQQLQARLMKE 22  
 CC  
 CC Sequence 17 AA;  
 CC Query Match 40.1%; Score 87; DB 5; Length 17;  
 CC Best Local Similarity 100.0%; Pred. No. 1.6e-05; Mismatches 0; Indels 0; Gaps 0;  
 CC Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC  
 Qy 29 WRLEPEDGTALCFIF 43  
 Db 1 WRLEPEDGTALCFIF 15  
 CC  
 CC RESULT 11  
 ID AAU85011  
 AC AAU85011 standard; peptide; 17 AA.  
 XX  
 AC AAU85011;  
 XX  
 DT 08-MAY-2002 (first entry)

RESULT 12  
 AAR67809  
 ID AAR67809 standard; peptide; 16 AA.  
 XX  
 AC AAR67809;  
 XX  
 DT

DT 25-MAR-2003 (revised)  
 DT 22-AUG-1995 (first entry)  
 XX BAGE tumor rejection antigen peptide.  
 DE BAGE; tumor rejection antigen precursor; diagnosis; HLA;  
 KW human leukocyte antigen MHC; major histocompatibility complex; TRAP;  
 KW cancer; melanoma.  
 XX Synthetic.  
 OS XX  
 PN WO9500159-A1.  
 XX  
 PD 05-JAN-1995.  
 XX  
 PF 10-JUN-1994; 94WO-US006534.  
 XX  
 PR 17-JUN-1993; 93US-00079110.  
 PR 15-FEB-1994; 94US-00196630.  
 XX  
 PA (LUDWIG) LUDWIG INST CANCER RES.  
 XX  
 PI Van Der Bruggen P, Boon-Faileur T, Coulie P, Renaud J;  
 XX WPI: 1995-051741/07.  
 XX  
 Nucleic acid coding for a tumour rejection antigen precursor - used to  
 PT develop prods. for the diagnosis and therapy of cancers, partic.  
 PT melanomas.  
 XX  
 PS Claim 21; Page 20; 33pp; English.  
 XX  
 CC This sequence encodes the tumor rejection antigen Peptide BAGE. The  
 CC peptide may be used in the diagnosis and therapy of cancers, e.g.  
 CC melanomas. (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 SQ Sequence 16 AA;  
 XX  
 Query Match 32.9%; Score 70; DB 2; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 0.0045;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 XX  
 RESULT 13  
 AY10635 ID AY10635 standard; peptide; 16 AA.  
 XX  
 AC AAY10635;  
 DT 12-MAY-1999 (first entry)  
 DE Peptide antigen SEQ ID NO:565.  
 XX  
 KW Cytotoxic T-lymphocyte response; CTL; antigen; lymphatic system;  
 KW immunisation; tumour; infectious disease; immunotherapy; cancer;  
 KW malignant melanoma; viral disease; hepatitis; AIDS.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 PN WO9902183-A2.  
 XX  
 PD 21-JAN-1999.  
 XX  
 PF 10-JUL-1998; 98WO-US014289.  
 XX  
 PR 10-JUL-1997; 97CA-02209815.  
 PR 10-DEC-1997; 97US-0098320.  
 XX  
 PA (CTLI-) CTL IMMUNOTHERAPIES CORP.  
 XX  
 PI Kuendig TM, Simard JJL;  
 DR WPI; 1999-120514/10.  
 XX  
 Inducing a cytotoxic T lymphocyte response - by maintaining a level of  
 PT antigen in the lymphatic system of a mammal so as to provide a sustained  
 PT CTL response, used to treat, e.g. AIDS.  
 XX  
 Disclosure; Page 52; 199DP; English.  
 XX  
 The present invention describes a method of inducing and/or sustaining an  
 CC immunological cytotoxic T lymphocyte (CTL) response in a mammal. The  
 CC method comprises: (a) delivering an antigen to the mammal at a level to  
 CC induce an immunological CTL response in the mammal; and (b) maintaining  
 CC the level of the antigen in the mammal's lymphatic system to maintain the  
 CC immunologic CTL response. The method can be used for the delivery of e.g.  
 CC a differentiation antigen, a tumour-specific multipeptide antigen, an  
 CC embryonic antigen, an oncogene antigen, a mutated tumour-suppressor gene  
 CC antigen, or a viral antigen. They can be used for the treatment of  
 CC disease such as cancer, e.g. malignant melanoma or infectious disease,  
 CC e.g. viral disease such as hepatitis or AIDS. Sustained antigen delivery  
 CC to the lymphatic system provides for potent CTL stimulation that takes  
 CC place in the milieu of the lymphoid organ, and it sustains stimulation  
 CC that is necessary to keep the CTL active, cytotoxic and recirculating through  
 CC the body. AAY10671 to AAY10639 represent examples of peptide antigens  
 CC given in the present invention  
 XX  
 SQ Sequence 16 AA;  
 XX  
 Query Match 32.9%; Score 70; DB 2; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 0.0045;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 XX  
 Qy 1 MAARAVFLAISQQLQ 16  
 Db 1 MAARAVFLAISQQLQ 16  
 XX  
 RESULT 14  
 ABG80316 ID ABG80316 standard; peptide; 16 AA.  
 XX  
 AC ABG80316;  
 XX  
 DT 29-AUG-2003 (revised)  
 DT 15-NOV-2002 (first entry)  
 DE MHC class I molecule, viral epitope #564.  
 XX  
 KW Major histocompatibility complex; MHC; MHC class I molecule; virus;  
 KW epitope; cytotoxic T lymphocyte response; CTL response; lymphatic system;  
 KW antigen; immunogenic; malignant tumour; carcinoma; melanoma; leukaemia;  
 KW lymphoma; infectious disease; hepatitis; malaria; measles; tuberculosis;  
 KW acquired immune deficiency syndrome; AIDS.  
 XX  
 Viruses.  
 XX  
 WO200262368-A2.  
 XX  
 PD 15-AUG-2002.  
 XX  
 PF 22-JAN-2002; 2002WO-US002033.  
 XX  
 PR 02-FEB-2001; 2001US-00776232.  
 XX  
 PA (CTLI-) CTL IMMUNOTHERAPIES CORP.  
 XX  
 PI Kuendig TM, Simard JJL;  
 DR WPI; 2002-657505/70.  
 XX

PT Inducing or sustaining immunological cytotoxic T lymphocyte response in a mammal, useful for treating a mammal with malignant tumor or infectious disease, by directly administering an antigen to the lymphatic system of the mammal.

PS Disclosure: Page 46; 73PP; English.

XX The invention relates to a method of inducing and/or sustaining an immunological cytotoxic T lymphocyte (CTL) response in a mammal comprising administering directly to the lymphatic system of the mammal: (a) an antigen in the form of a polypeptide; (b) a vector comprising a nucleic acid encoding the antigen; or (c) a non-peptide antigen. The method is useful for inducing and/or sustaining CTL response in a mammal. This is particularly useful for treating a mammal having a malignant tumour (e.g. carcinoma, melanoma, leukaemia or lymphoma) or infectious disease (e.g. hepatitis, acquired immune deficiency syndrome (AIDS), malaria, measles or tuberculosis), or in an animal having a predisposition to these diseases. The mammal may be dogs, cats, mice, cattle, sheep, pigs, goats, rabbits, or preferably humans. ABG79733-ABG0319 represent viral epitopes on major histocompatibility complex (MHC) class I molecules used in the method of the invention. (Updated on 29-AUG-2003 to standardise OS field)

XX Sequence 16 AA;

Query Match 32.9%; Score 70; DB 5; Length 16;  
Best Local Similarity 100.0%; Pred. No. 0.0045; Mismatches 0; Indels 0; Gaps 0;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MAARAVFLAQSQQI 16  
Db 1 MAARAVFLAQSQQI 16

RESULT 15

ID AAB48007 standard; protein; 384 AA.  
XX AAB48007;  
AC XX  
DT 19-MAR-2001 (First entry)  
XX Human Sphingosine kinase protein sequence.

DE Sphingosine kinase; antiarthritic; antiasthmatic; antiarteriosclerotic; antinflammatory; neuroprotective; antibacterial; immunosuppressive; human.  
XX Homo sapiens.  
OS Pitson SM, Wattenberg BW, Xia P, D'andrea RJ, Gamble JR;  
Vadas MA;

XX PN WO200070028-A1.  
XX DR 2001-016227/02.  
XX PD 23-NOV-2000.  
PF 12-MAY-2000; 2000WO-AU000457.

XX PR 13-MAY-1999; 99AU-00000339.  
PR 08-JUL-1999; 99AU-00001504.  
XX PA (JOHJ ) JOHNSON & JOHNSON RES PTY LTD.

XX PI Pitson SM, Wattenberg BW, Xia P, D'andrea RJ, Gamble JR;  
Vadas MA;  
XX DR WPI; 2001-016227/02.  
DR N-PSDB; AACB84161.  
XX Novel sphingosine kinase protein and nucleic acid molecules for diagnosis, prophylaxis and treatment of rheumatoid arthritis, asthma, atherosclerosis, inflammation, meningitis, multiple sclerosis and septic shock.

XX Claim 9; Fig 7a; 100pp; English.

XX This represents a human sphingosine kinase (SK) protein. The human SK protein, encoding nucleic acids and modulators are useful for modulating expression, functional activity or cellular functional activity of sphingosine kinase in a subject and also for treating a mammal by modulating the activity of SK. Diseases treated by regulating SK cellular activity include rheumatoid arthritis, asthma, atherosclerosis, inflammation, meningitis, multiple sclerosis and septic shock.

XX CC This represents a human sphingosine kinase (SK) protein. The human SK protein, encoding nucleic acids and modulators are useful for modulating expression, functional activity or cellular functional activity of sphingosine kinase in a subject and also for treating a mammal by modulating the activity of SK. Diseases treated by regulating SK cellular activity include rheumatoid arthritis, asthma, atherosclerosis, inflammation, meningitis, multiple sclerosis and septic shock.

CC Query Match 29.6%; Score 63; DB 4; Length 384;  
Best Local Similarity 37.2%; Pred. No. 1.8;  
Matches 16; Conservatice 9; Mismatches 12; Indels 6; Gaps 1;

CC Qy 3 ARAVFLAQSQQIQRMEES----PVVSWLEPEDGTAL 39  
Db 295 SRAMLRLFLAMEKGRHMEYECPYLVVYCPYLVYVPAFRLPEPDGKGM 337

CC Sequence 384 AA;

XX Search completed: July 7, 2004, 17:07:44  
Job time : 57 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: July 7, 2004, 17:06:43 ; Search time 23 Seconds  
(without alignments)  
96.518 Million cell updates/sec

Title: US-10-081-108-2

Perfect score: 213 MAARAVFLAASQQLIQARLM.....SPVVSWRLEPEDGFGALCFIF 43

Sequence: 1

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51623971 residues

Total number of hits satisfying chosen parameters:

389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*

1: /cgcn2\_6/ptodata/2/iaa/5A\_COMB\_pep:\*

2: /cgcn2\_6/ptodata/2/iaa/5B\_COMB\_pep:\*

3: /cgcn2\_6/ptodata/2/iaa/6A\_COMB\_pep:\*

4: /cgcn2\_6/ptodata/2/iaa/6B\_COMB\_pep:\*

5: /cgcn2\_6/ptodata/2/iaa/PCTUS\_COMB\_pep:\*

6: /cgcn2\_6/ptodata/2/iaa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	213	100.0	43	2	US-08-389-360-2	Sequence 2, Appli
2	213	100.0	43	3	US-08-389-328-2	Sequence 2, Appli
3	213	100.0	43	4	US-09-435-24-2	Sequence 2, Appli
4	213	100.0	43	4	/US-09-382-197-2	Sequence 2, Appli
5	98	46.0	22	1	US-08-196-630A-8	Sequence 8, Appli
6	70	32.9	16	1	US-08-196-630A-9	Sequence 9, Appli
7	62	29.1	54	4	US-09-621-576-5718	Sequence 328, AD
8	62	29.1	293	4	US-09-558-328	Sequence 328, AD
9	62	29.1	384	4	US-09-970-516-2	Sequence 2, Appli
10	57	26.8	398	2	US-08-284-391B-39	Sequence 29, Appli
11	57	26.8	398	3	US-09-218-550-29	Sequence 29, Appli
12	57	26.8	433	2	US-08-867-149-1	Sequence 1, Appli
13	57	26.8	433	2	US-08-374-41	Sequence 1, Appli
14	57	26.8	433	3	US-09-100-109A-1	Sequence 1, Appli
15	57	26.8	458	3	US-08-466-368-4	Sequence 4, Appli
16	57	26.8	462	2	US-08-417-495-5	Sequence 5, Appli
17	57	26.8	462	2	US-08-284-391B-5	Sequence 5, Appli
18	57	26.8	462	3	US-09-218-950-5	Sequence 5, Appli
19	57	26.8	462	5	PCT-US92-01785-6	Sequence 5, Appli
20	57	26.8	462	5	PCT-US95-00454-6	Sequence 4, Appli
21	57	26.8	532	2	US-08-417-495-6	Sequence 6, Appli
22	57	26.8	532	2	US-08-417-491B-6	Sequence 6, Appli
23	57	26.8	532	3	US-09-218-550-6	Sequence 6, Appli
24	57	26.8	532	5	PCT-US92-01785-6	Sequence 6, Appli
25	57	26.8	532	5	PCT-US95-00454-6	Sequence 6, Appli
26	57	26.8	575	2	US-08-417-495-4	Sequence 0; Gaps
27	57	26.8	575	2	US-08-284-391B-4	Sequence 4, Appli

## ALIGNMENTS

RESULT 1  
US-08-389-360-2  
; Sequence 2, Appli US/08389360  
; Patent No. 5877017  
; GENERAL INFORMATION:  
; APPLICANT: van der Bruggen et al.  
; TITLE OF INVENTION: ISOLATED PEPTIDE WHICH FORMS COMPLEXES WITH MHC MOLECULE HLA-Cw\*1601 AND USES  
; TITLE OF INVENTION: THEREOF  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Felipe & Lynch  
; STREET: 805 Third Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10022  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: Wordperfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/389,360  
; CLASSIFICATION: 436  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: 08/196,630  
; FILING DATE: February 15, 1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/079,110  
; FILING DATE: June 17, 1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pasqualini, Patricia A.  
; REGISTRATION NUMBER: 34,894  
; REFERENCE DOCKET NUMBER: LOD 5310.2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 688-9200  
; TELEFAX: (212) 831-3884  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 43  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-08-389-360-2

Query Match 100.0%; Score 213, DB 2;  
Best Local Similarity 100.0%; Pred. No. 1.1e-26;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAARAVFLAASQIQLQARLMKEESPVVSVRLEPDGTALCFIF 43  
 Db 1 MAARAVFLAASQIQLQARLMKEESPVVSVRLEPDGTALCFIF 43

RESULT 2  
 US-09-038-328-2  
 ; Sequence 2, Application US/09038328

; Patent No. 6110694

GENERAL INFORMATION:

APPLICANT: van der Bruggen et al.

TITLE OF INVENTION: ISOLATED PEPTIDE WHICH FORMS COMPLEXES

WITH MHC MOLECULE HLA-Cw\*1601 AND USES

TITLE OF INVENTION: THEREOF

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: Felfe & Lynch

STREET: 805 Third Avenue

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10022

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette

COMPUTER: IBM PS/2

OPERATING SYSTEM: PC-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/435,524

FILING DATE: 08-Nov-6465184-1999

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/038,328

FILING DATE: <Unknown>

APPLICATION NUMBER: 08/079,110

FILING DATE: Jun 17, 1993

ATTORNEY/AGENT INFORMATION:

NAME: Pasqualini, Patricia A.

REGISTRATION NUMBER: 34,894

REFERENCE/DOCKET NUMBER: LUD 5310.2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 688-9200

TELEFAX: (212) 838-3884

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 43

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-435,524-2

RESULT 4  
 US-09-382-497-2

; Sequence 2, Application US/09382497

; Patent No. 6638512

GENERAL INFORMATION:

APPLICANT: van der Bruggen et al.

TITLE OF INVENTION: ISOLATED PEPTIDE WHICH FORMS COMPLEXES

WITH MHC MOLECULE HLA-Cw\*1601 AND USES

TITLE OF INVENTION: THEREOF

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: Felfe & Lynch

STREET: 805 Third Avenue

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10022

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette

COMPUTER: IBM PS/2

OPERATING SYSTEM: PC-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/382,497

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/389,360

CLASSIFICATION:

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: Felfe & Lynch  
 STREET: 805 Third Avenue  
 CITY: New York  
 STATE: New York  
 COUNTRY: USA

ZIP: 10022

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette

COMPUTER: IBM PS/2

OPERATING SYSTEM: PC-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/435,524

FILING DATE: 08-Nov-6465184-1999

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/038,328

FILING DATE:

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-038-328-2

RESULT 3  
 US-09-435-524-2

; Sequence 2, Application US/09435524

; Patent No. 6638514

GENERAL INFORMATION:

APPLICANT: van der Bruggen et al.

TITLE OF INVENTION: ISOLATED PEPTIDE WHICH FORMS COMPLEXES

WITH MHC MOLECULE HLA-Cw\*1601 AND USES

TITLE OF INVENTION: THEREOF

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: Felfe & Lynch

STREET: 805 Third Avenue

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10022

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette

COMPUTER: IBM PS/2

OPERATING SYSTEM: PC-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/382,497

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/389,360

CLASSIFICATION:

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

FILING DATE:  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: 08/079,110  
 FILING DATE: June 17, 1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Pasqualini, Patricia A.  
 REGISTRATION NUMBER: 34,894  
 TELEPHONE: (212) 688-9200  
 TELEFAX: (212) 838-3884  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 43  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: Linear  
 US-09-312-497-2

Query Match 100.0%; Score 213; DB 4; Length 43;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-26;  
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAARAVFLAQLQQLQARLMLKE 43  
 Db 1 MAARAVFLAQLQQLQARLMLKE 43

RESULT 5  
 US-08-196-630A-8  
 / Sequence 8, Application US/08196630A  
 / Patent No. 5683866  
 / GENERAL INFORMATION:  
 / APPLICANT: van der Bruggen, Pierre  
 / APPLICANT: Boon-Falleur, Thierry  
 / TITLE OF INVENTION: ISOLATED PEPTIDES WHICH FORM  
 / COMPLEXES WITH MHC MOLECULE HLA-C-CLONE 10 AND USES  
 / TITLE OF INVENTION: THEREOF  
 / NUMBER OF SEQUENCES: 10  
 / CORRESPONDENCE ADDRESS:  
 / ADDRESSEE: Felfe & Lynch  
 / STREET: 805 Third Avenue  
 / CITY: New York City  
 / STATE: New York  
 / ZIP: 10022  
 / COMPUTER READABLE FORM:  
 / MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage  
 / COMPUTER: IBM PS/2  
 / OPERATING SYSTEM: PC-DOS  
 / SOFTWARE: Wordperfect  
 / CURRENT APPLICATION NUMBER: US/08/196,630A  
 / APPLICATION NUMBER: 08/079,110  
 / FILING DATE: 15-FEB-1994  
 / CLASSIFICATION: 436  
 / PRIOR APPLICATION DATA:  
 / APPLICATION NUMBER: 08/079,110  
 / FILING DATE: 17-JUN-1993  
 / ATTORNEY/AGENT INFORMATION:  
 / NAME: Hanson, No. 5683866man D.  
 / REGISTRATION NUMBER: 30,946  
 / REFERENCE/DOCKET NUMBER: LUD 5310.1  
 / TELECOMMUNICATION INFORMATION:  
 / TELEPHONE: (212) 688-9200  
 / TELEFAX: (212) 838-3884  
 / INFORMATION FOR SEQ ID NO: 9:  
 / SEQUENCE CHARACTERISTICS:  
 / LENGTH: 16 amino acid residues  
 / TYPE: amino acid residues  
 / TOPOLOGY: linear  
 US-08-196-630A-9;

Query Match 32.9%; Score 70; DB 1; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 0.00016;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAARAVFLAQLQQLQARLMLKE 16  
 Db 1 MAARAVFLAQLQQLQARLMLKE 16

RESULT 7  
 US-09-631-976-5718  
 / Sequence 5718, Application US/09621976  
 / Patent No. 6639063  
 / GENERAL INFORMATION:  
 / APPLICANT: Giardano, J.Y.  
 / APPLICANT: Dumas Milne Edwards, J.B.  
 / TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
 / FILE REFERENCE: GENSET\_054PR2  
 / CURRENT APPLICATION NUMBER: US/09/621,976  
 / CURRENT FILING DATE: 2000-07-21

Query Match 46.0%; Score 98; DB 1; Length 22;

NUMBER OF SEQ ID NOS: 19335  
 SOFTWARE: Patent .pm.  
 SEQ ID NO: 5718  
 LENGTH: 54  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 FEATURE: SIGNAL  
 NAME/KEY: SIGNAL  
 LOCATION: -50..-1  
 US-09-621-976-5718

Query Match 29.1%; Score 62; DB 4; Length 54;  
 Best Local Similarity 66.7%; Pred. No. 0.015; N mismatches 6; Indels 0; Gaps 0;

```

Qy  22 BESPVVSWRLEPDGTL 39
Db   22 EESSVARWRLPEDGMVL 39

```

RESULT 8  
 US-09-205-258-328  
 ; Sequence 328, Application US/09205258  
 ; Patent No. 6525174.

GENERAL INFORMATION:

APPLICANT: Young et al.  
 TITLE OF INVENTION: 207 Human Secreted Proteins  
 FILE REFERENCE: P2007P1

CURRENT APPLICATION NUMBER: US/09/205,258  
 CURRENT FILING DATE: 1998-12-04  
 EARLIER APPLICATION NUMBER: PCT/US98/11422  
 EARLIER FILING DATE: 1998-06-04  
 EARLIER APPLICATION NUMBER: 60/048,885  
 EARLIER FILING DATE: 1997-06-06  
 EARLIER APPLICATION NUMBER: 60/049,375  
 EARLIER FILING DATE: 1997-06-06  
 EARLIER APPLICATION NUMBER: 60/048,881  
 EARLIER FILING DATE: 1997-06-06  
 EARLIER APPLICATION NUMBER: 60/048,880  
 EARLIER FILING DATE: 1997-06-06  
 EARLIER APPLICATION NUMBER: 60/048,896  
 EARLIER FILING DATE: 1997-06-06  
 EARLIER APPLICATION NUMBER: 60/049,020  
 EARLIER FILING DATE: 1997-06-06  
 EARLIER APPLICATION NUMBER: 60/048,894  
 EARLIER FILING DATE: 1997-06-06  
 EARLIER APPLICATION NUMBER: 60/048,971  
 EARLIER FILING DATE: 1997-06-06  
 EARLIER APPLICATION NUMBER: 60/048,964  
 EARLIER FILING DATE: 1997-06-06  
 EARLIER APPLICATION NUMBER: 60/048,901  
 EARLIER FILING DATE: 1997-06-06  
 EARLIER APPLICATION NUMBER: 60/048,892  
 EARLIER FILING DATE: 1997-06-06  
 EARLIER APPLICATION NUMBER: 60/048,949  
 EARLIER FILING DATE: 1997-06-06  
 EARLIER APPLICATION NUMBER: 60/048,875  
 EARLIER FILING DATE: 1997-06-06  
 EARLIER APPLICATION NUMBER: 60/049,374  
 EARLIER FILING DATE: 1997-06-06  
 EARLIER APPLICATION NUMBER: 60/048,883  
 EARLIER FILING DATE: 1997-06-06  
 EARLIER APPLICATION NUMBER: 60/048,897  
 EARLIER FILING DATE: 1997-06-06  
 EARLIER APPLICATION NUMBER: 60/048,898  
 EARLIER FILING DATE: 1997-06-06  
 EARLIER APPLICATION NUMBER: 60/048,962  
 EARLIER FILING DATE: 1997-06-06  
 EARLIER APPLICATION NUMBER: 60/048,963  
 EARLIER FILING DATE: 1997-06-06  
 EARLIER APPLICATION NUMBER: 60/048,877  
 EARLIER FILING DATE: 1997-06-06  
 EARLIER APPLICATION NUMBER: 60/048,878  
 EARLIER FILING DATE: 1997-06-06  
 EARLIER APPLICATION NUMBER: 60/070,923  
 EARLIER FILING DATE: 1997-12-18  
 EARLIER APPLICATION NUMBER: 60/092,921  
 EARLIER FILING DATE: 1998-07-15  
 EARLIER APPLICATION NUMBER: 60/094,657  
 EARLIER FILING DATE: 1998-07-30  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO 328  
 LENGTH: 293  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: SITE  
 LOCATION: (36)  
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 US-09-205-258-328

Query Match 29.1%; Score 62; DB 4; Length 293;  
 Best Local Similarity 40.0%; Prod. No. 0.14;  
 Matches 16; Conservative 8; Mismatches 10; Indels 6; Gaps 1;

Qy 3 ARAVFLAASQIQLQARLMKES-----PVVSWRLPEDG 36
 Db 3 ARAVFLAASQIQLQARLMKES-----PVVSWRLPEDG 36
 Db 204 SRAMLLRLFLAMEKGRRMVEYECPYLVYPVVAFRLEPKDG 243

RESULT 9  
 US-09-970-516-2  
 Sequence 2, Application US/09970516  
 ; Patent No. 6610534  
 ; GENERAL INFORMATION:  
 ; APPLICANT: No. 6610534 artis AG  
 ; TITLE OF INVENTION: Induction of blood vessel formation through administration of  
 ; polynucleotides encoding sphingosine kinases  
 ; FILE REFERENCE: 4-16167  
 ; CURRENT APPLICATION NUMBER: US/09/970,516  
 ; CURRENT FILING DATE: 2001-10-04  
 ; NUMBER OF SEQ ID NOS: 6  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 2  
 ; LENGTH: 384

;

TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-970-516-2

Query Match 29.1%; Score 62; DB 4; Length 384;  
Best Local Similarity 40.0%; Pred. No. 0.19;  
Matches 16; Conservative 8; Mismatches 10; Indels 6; Gaps 1;

QY 3 ARAVFLAQLQQLQARLMEES----PVVSWRLEPENG 36  
Db 295 SRAMVLRLFAMEKGRHMYECPIVYVWAFREPKD 334

RESULT 10  
US-08-284-391B-29  
Sequence 29, Application US/08284391B  
Patent No. 5851828

GENERAL INFORMATION:  
APPLICANT: Seed, Brian  
APPLICANT: Banapour, Babak  
APPLICANT: Romeo, Charles  
APPLICANT: Kolanus, Waldemar  
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED CELLS BY CHIMERIC CD4 RECEPTOR- BEARING CELLS

NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Clark & Elbing LLP  
STREET: 176 Federal Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ FOR Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/284,391B  
FILING DATE: 02-AUG-1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/195,395  
FILING DATE: 14-FEB-1994  
APPLICATION NUMBER: 07/847,566  
FILING DATE: 06-MAR-1992  
APPLICATION NUMBER: 07/665,961  
FILING DATE: 07-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Elbing, Karen L.  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/247001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045

INFORMATION FOR SEQ ID NO: 29:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 398 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-218-950-29

RESULT 11  
US-09-218-950-29  
Sequence 29, Application US/09218950  
Patent No. 6294240

GENERAL INFORMATION:  
APPLICANT: Seed, Brian  
APPLICANT: Banapour, Babak  
APPLICANT: Romeo, Charles  
APPLICANT: Kolanus, Waldemar  
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED CELLS BY CHIMERIC CD4 RECEPTOR- BEARING CELLS

NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Clark & Elbing LLP  
STREET: 176 Federal Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ FOR Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/218,950  
FILING DATE: 02-AUG-1994  
CLASSIFICATION:  
PRIOR APPLICATION NUMBER: 08/284,391  
FILING DATE: 02-AUG-1994  
APPLICATION NUMBER: 08/195,395  
FILING DATE: 14-FEB-1994  
APPLICATION NUMBER: 07/847,566  
FILING DATE: 06-MAR-1992  
APPLICATION NUMBER: 07/665,961  
FILING DATE: 07-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Elbing, Karen L.  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/247001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045

INFORMATION FOR SEQ ID NO: 29:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 398 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-218-950-29

Query Match 26.8%; Score 57; DB 3; Length 398;  
Best Local Similarity 35.1%; Pred. No. 1.3;  
Matches 13; Conservative 7; Mismatches 15; Indels 2; Gaps 1;

QY 6 VFLALSAQLQQLQARLMEESPVVSWRLEPENGTCI 42  
Db 338 IMLSLXLENKEAKVSKREKPV-WVNPEAGNMWQCHL 372

RESULT 12  
US-08-067-149-1  
Sequence 1, Application US/08867149  
Patent No. 5912176

GENERAL INFORMATION:  
APPLICANT: Wang, Chang Yi  
TITLE OF INVENTION: Antibody Against a Host Cell  
TITLE OF INVENTION: Antigen Complex for Pre- and Post-Exposure Protection From Infection by HIV Primary Isolates

NUMBER OF SEQUENCES: 2  
 ADDRESSEE: Maria C.H. Lin  
 STREET: 345 Park Avenue  
 CITY: New York  
 STATE: NY  
 COUNTRY: USA  
 ZIP: 10154-0053  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: FLOPPY disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/867,149  
 FILING DATE: 28-Feb-1997  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/657,149  
 FILING DATE: 03-June-1997  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Maria C.H. Lin  
 REGISTRATION NUMBER: 29,323  
 REFERENCE/DOCKET NUMBER: 1151-4145  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 751-6849  
 TELEFAX: (212) 751-6849  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 433 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-08-867-149-1

Query Match Score 57; DB 2; Length 433;  
 Best Local Similarity 35.1%; Pred. No. 1.4;  
 Matches 13; Conservative 7; Mismatches 15; Indels 2; Gaps 1;

Qy 6 VPIALSAQLQARLMKEESPVVSVRLEPDGTALCFI 42  
 Db 313 LMLSLKLENKEAKVSKREKPV-WVLNPEAGMNGQCLL 347

RESULT 14  
 US-09-100-409A-1  
 Sequence 1, Application US/09100409A  
 Patent No. 6090388  
 GENERAL INFORMATION:  
 APPLICANT: Wang, Chang Yi  
 TITLE OF INVENTION: PEPTIDE COMPOSITION FOR PREVENTION AND TREATMENT OF HIV INFECTION AND IMMUNE DISORDERS  
 TITLE OF INVENTION: IMMUNE DISORDERS  
 NUMBER OF SEQUENCES: 64  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: MORGAN & FINNEGAN  
 STREET: 345 Park Avenue  
 CITY: New York  
 STATE: NY  
 COUNTRY: USA  
 ZIP: 10154-0054  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/100,409A  
 FILING DATE:  
 CLASSIFICATION: 514  
 ATTORNEY/AGENT INFORMATION:  
 NAME:  
 REGISTRATION NUMBER: 1151-4154  
 REFERENCE/DOCKET NUMBER: 1151-4154  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 212-758-4800  
 TELEFAX: 212-751-6849  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 433 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-09-100-409A-1

Query Match Score 57; DB 3; Length 433;  
 Best Local Similarity 35.1%; Pred. No. 1.4;  
 Matches 13; Conservative 7; Mismatches 15; Indels 2; Gaps 1;

Qy 6 VPIALSAQLQARLMKEESPVVSVRLEPDGTALCFI 42  
 Db 313 LMLSLKLENKEAKVSKREKPV-WVLNPEAGMNGQCLL 347

RESULT 15  
 US-08-867-149-1  
 Sequence 1, Application US/08808374  
 Patent No. 5961976  
 GENERAL INFORMATION:  
 APPLICANT: Wang, Chang Yi  
 TITLE OF INVENTION: Antibody Against a Host Cell  
 TITLE OF INVENTION: Antigen Complex for Pre- and Post-Exposure  
 TITLE OF INVENTION: Protection from Infection by HIV Primary Isolates  
 NUMBER OF SEQUENCES: 2  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Maria C.H. Lin  
 STREET: 345 Park Avenue  
 CITY: New York  
 STATE: NY  
 COUNTRY: USA  
 ZIP: 10154-0053  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/808,374  
 FILING DATE: 28-Feb-1997  
 CLASSIFICATION: 435

Qy 6 VFLALSAQLQARLMKEESPVVSRLPEDGTALCFI 42  
 Db : |:| : |:| : |:| : |:| : |:| : |:| : |:| : |:|  
 313 LMLSLKLENKEAKVSKREKEV- -WVLNPEAGMNCCLL 347

## RESULT 15

US-08-466-368-4

Sequence 4, Application US/08466368

## GENERAL INFORMATION:

APPLICANT: Maddon, Paul J.  
 Littman, Dan R.APPLICANT: Chess, Leonard  
 Axel, RichardAPPLICANT: Weis, Robin  
 McDougal, J. S.TITLE OF INVENTION: DNA ENCODING THE T CELL SURFACE PROTEIN  
 TITLE OF INVENTION: T4 AND USE OF FRAGMENTS OF T4 IN THE TREATMENT OF AIDS

NUMBER OF SEQUENCES: 21

## CORRESPONDENCE ADDRESS:

ADDRESSEE: Cooper S. Dunham LLP  
 STREET: 1185 Avenue of Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

## CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/466,368

FILING DATE: 06-JUN-1995

CLASSIFICATION: 435

## ATTORNEY/AGENT INFORMATION:

NAME: White, John P.

REGISTRATION NUMBER: 28,678

REFERENCE/DOCKET NUMBER: 24577-E1-B/JPW/AKC

## TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-278-0400

TELEFAX: 212-391-0525

## INFORMATION FOR SEQ ID NO: 4:

## SEQUENCE CHARACTERISTICS:

LENGTH: 458 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-466-368-4

Query Match 26.8%; Score 57; DB 3; Length 458;  
 Best Local Similarity 35.1%; Pred. No. 1.5%;  
 Matches 13; Conservative 7; Mismatches 15; Indels 2; Gaps 1;Qy 6 VFLALSAQLQARLMKEESPVVSRLPEDGTALCFI 42  
 Db : |:| : |:| : |:| : |:| : |:| : |:| : |:|  
 338 LMLSLKLENKEAKVSKREKEV- -WVLNPEAGMNCCLL 372

Search completed: July 7, 2004, 17:10:12

Job time : 28 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 Compugen Ltd.

## CM protein - Protein search, using sw model

Run on: July 7, 2004, 17:09:14 ; Search time 46 Seconds  
 (without alignments)

290.983 Million cell updates/sec

Title: US-10-081-108-2

Perfect score: 213

Sequence: 1 MARARVFLAQLSQQLQARLM.....SPVVSWRLEPEDGDTALCFIF 43

Scoring table: BLOSUM62

Gapp 10.0 , Gapext 0.5

Searched: 1276540 seqs, 311283816 residues

Total number of hits satisfying chosen parameters: 1276540

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:<sup>\*</sup>

1: /cgmn\_6/ptodata/1/pubpaas/US07\_PUBCOMB.pep:  
 2: /cgmn\_6/ptodata/1/pubpaas/PCT\_NEW\_PUB\_pep:  
 3: /cgmn\_6/ptodata/1/pubpaas/US06\_PUB\_pep:  
 4: /cgmn\_6/ptodata/1/pubpaas/US05\_PUBCOMB.pep:  
 5: /cgmn\_6/ptodata/1/pubpaas/US07\_NEW\_PUB\_pep:  
 6: /cgmn\_6/ptodata/1/pubpaas/PCTRS\_PUBCOMB.pep:  
 7: /cgmn\_6/ptodata/1/pubpaas/US08\_NEW\_PUB\_pep:  
 8: /cgmn\_6/ptodata/1/pubpaas/US08\_PUBCOMB.pep:  
 9: /cgmn\_6/ptodata/1/pubpaas/US09\_PUBCOMB.pep:  
 10: /cgmn\_6/ptodata/1/pubpaas/US09B\_PUBCOMB.pep:  
 11: /cgmn\_6/ptodata/1/pubpaas/US09C\_PUBCOMB.pep:  
 12: /cnr2\_6/ptodata/1/pubpaas/US09\_NEW\_PUB\_pep:  
 13: /cnr2\_6/ptodata/1/pubpaas/US09\_PUBCOMB.pep:  
 14: /cnr2\_6/ptodata/1/pubpaas/US10\_PUBCOMB.pep:  
 15: /cnr2\_6/ptodata/1/pubpaas/US10C\_PUBCOMB.pep:  
 16: /cnr2\_6/ptodata/1/pubpaas/US10I\_NEW\_PUB\_pep:  
 17: /cnr2\_6/ptodata/1/pubpaas/US06\_NEW\_PUB\_pep:  
 18: /cnr2\_6/ptodata/1/pubpaas/US06\_PUBCOMB.pep:  
 RESULT 1  
 US-10-296-734-825

; Sequence 825, Application US-10296734  
 ; Publication No. US20040054137A1

; GENERAL INFORMATION:

; APPLICANT: Thompson, Scott A

; TITLE OF INVENTION: Synthetic molecules and uses therefor

; FILE REFERENCE: Savine

; CURRENT APPLICATION NUMBER: US10/296,734

; PRIORITY DATE: 2003-08-04

; PRIOR APPLICATION NUMBER: AU PQ7761/00

; PRIORITY FILING DATE: 2000-05-26

; NUMBER OF SEQ ID NOS: 1507

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO: 825

; LENGTH: 43

; TYPE: PRT

; FEATURE:

; OTHER INFORMATION: BAGE consensus polypeptide

US-10-296-734-825

; Query Match 100.0%; Score 213; DB 12; Length 43;

; Best Local Similarity 100.0%; Pred. No. 6.7e-23;

; Matches 43; Conservative 0; Mismatches 0; Indexes 0; Gaps 0;

; ORGANISM: Artificial

; FEATURES:

; SUMMARYS

Result No.	Score	Query Match	Length	DB ID	Description
1	213	100.0	43	12 US-10-296-734-825	Sequence 825, App
2	213	100.0	43	14 US-10-207-655-01	Sequence 2, App
3	213	100.0	43	14 US-10-081-108-02	Sequence 2, App
4	156	73.2	30	12 US-10-296-734-1214	Sequence 1214, App
5	156	73.2	3541	12 US-10-296-734-1454	Sequence 1454, App
6	126	59.2	30	12 US-10-296-734-1212	Sequence 1212, App
7	87	40.8	17	12 US-10-296-734-1216	Sequence 1216, App
8	62	29.1	293	10 US-09-933-767-328	Sequence 328, App
9	62	29.1	293	12 US-10-004-860-528	Sequence 328, App
10	62	29.1	293	14 US-10-282-282-328	Sequence 288, App
11	62	333	15	US-10-264-237-585	Sequence 288, App
12	62	29.1	368	14 US-10-053-510-21	Sequence 21, App
13	62	29.1	368	15 US-10-348-502-21	Sequence 21, App
14	62	29.1	384	9 US-09-784-810A-2	Sequence 2, App
15	62	29.1	384	9 US-09-970-516-2	Sequence 2, App

; RESULT 2  
 US-10-207-655-81

; Sequence 81, Application US-10207655

; Publication No. US20030118592A1

; GENERAL INFORMATION:

; APPLICANT: Ledbetter, Jeffrey A.

APPLICANT: Hayden-Ledbetter, Martha S.  
 TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS  
 FILE REFERENCE: 39069.401C1  
 CURRENT APPLICATION NUMBER: US/10/207,655  
 CURRENT FILING DATE: 2003-07-25  
 NUMBER OF SEQ ID NOS: 426  
 SOFTWARE: PatentIn version 3.0  
 SEQ ID NO: 81  
 LENGTH: 43  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-10-207-655-81

Query Match 100.0%; Score 213; DB 14; Length 43;  
 Best Local Similarity 100.0%; Pred. No. 6.7e-23;  
 Matches 43; Conservative 0; Mismatches 0;  
 Indels 0; Gaps 0;

Qy 1 MAARAVFLAASQQLQARLMKEESPVVSRLEPDGTALCFIF 43  
 Db 1 MAARAVFLAASQQLQARLMKEESPVVSRLEPDGTALCFIF 43

RESULT 3  
 US-10-081-108-2  
 Sequence 2, Application US/10081108  
 Publication No. US20030138854A1  
 GENERAL INFORMATION:  
 APPLICANT: van der Bruggen et al.  
 TITLE OF INVENTION: ISOLATED PEPTIDE WHICH FORMS COMPLEXES WITH MHC MOLECULE HLA-Cw\*1601 AND USES THEREOF  
 NUMBER OF SEQUENCES: 9  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Feilje & Lynch  
 STREET: 805 Third Avenue  
 CITY: New York  
 STATE: New York  
 COUNTRY: USA  
 ZIP: 10022

COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette  
 COMPUTER SYSTEM: IBM PS/2  
 OPERATING SYSTEM: PC-DOS  
 SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/10/081,108  
 FILING DATE: 20-Feb-2002  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 09/435,524  
 FILING DATE: 08-Nov-2000  
 APPLICATION NUMBER: 09/038,328  
 FILING DATE: <Unknown>  
 APPLICATION NUMBER: 08/079,110  
 FILING DATE: June 17, 1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Pasqualini, Patricia A.  
 REGISTRATION NUMBER: 34,894  
 REFERENCE/DOCKET NUMBER: LUD 5310.2  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 638-9200  
 TELEX/FAX: (212) 838-3884  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 43  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Query Match 100.0%; Score 213; DB 14; Length 43;  
 Best Local Similarity 100.0%; Pred. No. 6.7e-23;

Matches 43; Conservative 0; Mismatches 0;  
 Indels 0; Gaps 0;  
 Qy 1 MAARAVFLAASQQLQARLMKEESPVVSRLEPDGTALCFIF 43  
 Db 1 MAARAVFLAASQQLQARLMKEESPVVSRLEPDGTALCFIF 43

RESULT 4  
 US-10-296-734-1214  
 Sequence 1214, Application US/10296734  
 Publication No. US20040054137A1  
 GENERAL INFORMATION:  
 APPLICANT: Thompson, Scott A.  
 TITLE OF INVENTION: Synthetic molecules and uses therefor  
 FILE REFERENCE: Savine  
 CURRENT APPLICATION NUMBER: US/10/296,734  
 CURRENT FILING DATE: 2003-08-04  
 PRIOR APPLICATION NUMBER: AU P27761/00  
 PRIOR FILING DATE: 2000-05-26  
 NUMBER OF SEQ ID NOS: 1507  
 SOFTWARE: PatentIn version 3.2  
 SEQ ID NO: 1214  
 LENGTH: 30  
 TYPE: PRT  
 ORGANISM: Artificial  
 FEATURE:  
 OTHER INFORMATION: BAGE segment 2

Query Match 73.2%; Score 156; DB 12;  
 Best Local Similarity 100.0%; Pred. No. 6.3e-15;  
 Matches 30; Conservative 0; Mismatches 0;  
 Indels 0; Gaps 0;

Qy 14 LIQARLMKEESPVVSRLEPDGTALCFIF 43  
 Db 1 LIQARLMKEESPVVSRLEPDGTALCFIF 30

RESULT 5  
 US-10-296-734-1454  
 Sequence 1454, Application US/10296734  
 Publication No. US20040054137A1  
 GENERAL INFORMATION:  
 APPLICANT: Thompson, Scott A.  
 TITLE OF INVENTION: Synthetic molecules and uses therefor  
 FILE REFERENCE: Savine  
 CURRENT APPLICATION NUMBER: US/10/296,734  
 CURRENT FILING DATE: 2003-08-04  
 PRIOR APPLICATION NUMBER: AU P27761/00  
 PRIOR FILING DATE: 2000-05-26  
 NUMBER OF SEQ ID NOS: 1507  
 SOFTWARE: PatentIn version 3.2  
 SEQ ID NO: 1454  
 LENGTH: 3541  
 TYPE: PRT  
 ORGANISM: Artificial  
 FEATURE:  
 OTHER INFORMATION: Melanoma cancer specific savine

Query Match 73.2%; Score 156; DB 12;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-12;  
 Matches 30; Conservative 0; Mismatches 0;  
 Indels 0; Gaps 0;

Qy 14 LIQARLMKEESPVVSRLEPDGTALCFIF 43  
 Db 2777 LIQARLMKEESPVVSRLEPDGTALCFIF 2806

RESULT 6  
 US-10-296-734-1212





; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,916  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/049,373  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,875  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/049,374  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,917  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,949  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,974  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,883  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,897  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,898  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,962  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,963  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,877  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/048,878  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/070,923  
 ; EARLIER FILING DATE: 1997-12-18  
 ; EARLIER APPLICATION NUMBER: 60/092,921  
 ; EARLIER FILING DATE: 1998-07-15  
 ; EARLIER APPLICATION NUMBER: 60/094,657  
 ; EARLIER FILING DATE: 1998-07-30  
 ; NUMBER OF SEQ ID NOS: 1227  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO: 328  
 ; LENGTH: 293  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: SITE  
 ; LOCATION: (36)  
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

US-10-023-282-328

Query Match 29.1%; Score 62; DB 14; Length 293;  
 Best Local Similarity 40.0%; Pred. No. 2.5;  
 Matches 16; Conservative 8; Mismatches 10; Indels 6; Gaps 1;

Qy 3 ARAVFLASQQLQARLMKKEES----PVYWSWRLEPEDG 36  
 Db 204 SRAMLLRFLAMEKGRRHMEYCPCPYLVYPVVAFRLEPKDG 243

RESULT 11

US-10-264-237-2585  
 ; Sequence 2585, Application US/10264237  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Birse et al.  
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
 ; CURRENT APPLICATION NUMBER: US/10/264,237  
 ; CURRENT FILING DATE: 2002-10-04  
 ; PRIOR APPLICATION NUMBER: PCT/US01/16450  
 ; PRIOR FILING DATE: 2001-05-18  
 ; PRIOR APPLICATION NUMBER: US 60/205,515  
 ; PRIOR FILING DATE: 2000-05-19  
 ; NUMBER OF SEQ ID NOS: 2876  
 ; SOFTWARE: PatentIn Ver. 3.1  
 ; SEQ ID NO: 2585

Query Match 29.1%; Score 62; DB 15; Length 368;  
 Best Local Similarity 40.0%; Pred. No. 3.3;  
 Matches 16; Conservative 8; Mismatches 10; Indels 6; Gaps 1;

Qy 3 ARAVFLASQQLQARLMKKEES----PVYWSWRLEPEDG 36  
 Db 279 SRAMLLRFLAMEKGRRHMEYCPCPYLVYPVVAFRLEPKDG 318

RESULT 13

US-10-348-052-21  
 ; Sequence 21, Application US/10348052  
 ; Publication No. US200301978A1

; GENERAL INFORMATION:  
 ; APPLICANT: Saba, Julie D.  
 ; TITLE OF INVENTION: SPHINGOSINE-1-PHOSPHATE LYASE POLYPEPTIDES,  
 ; TITLE OF INVENTION: POLYNUCLEOTIDES AND MODULATING AGENTS AND  
 ; TITLE OF INVENTION: METHODS OF USE THEREFOR  
 ; FILE REFERENCE: 200116-402C2  
 ; CURRENT APPLICATION NUMBER: US/10/053,510  
 ; CURRENT FILING DATE: 2002-01-17  
 ; NUMBER OF SEQ ID NOS: 21  
 ; SOFTWARE: Fast-SEQ for Windows Version 4.0  
 ; SEQ ID NO: 21  
 ; LENGTH: 368  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens

US-10-053-510-21

Query Match 29.1%; Score 62; DB 14; Length 368;  
 Best Local Similarity 40.0%; Pred. No. 3.3;  
 Matches 16; Conservative 8; Mismatches 10; Indels 6; Gaps 1;

Qy 3 ARAVFLASQQLQARLMKKEES----PVYWSWRLEPEDG 36  
 Db 279 SRAMLLRFLAMEKGRRHMEYCPCPYLVYPVVAFRLEPKDG 318

RESULT 15

US-10-348-052-21  
 ; Sequence 21, Application US/10348052  
 ; Publication No. US200301978A1

; GENERAL INFORMATION:  
 ; APPLICANT: Saba, Julie D.  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE MODULATION  
 ; TITLE OF INVENTION: OF SPHINGOLIPID METABOLISM AND/OR SIGNALING  
 ; FILE REFERENCE: 200116-405  
 ; CURRENT APPLICATION NUMBER: US/10/348,052  
 ; CURRENT FILING DATE: 2003-01-17  
 ; NUMBER OF SEQ ID NOS: 29  
 ; SOFTWARE: Fast-SEQ for Windows Version 4.0  
 ; SEQ ID NO: 21  
 ; LENGTH: 368  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens

US-10-348-052-21

Query Match 29.1%; Score 62; DB 15; Length 368;  
 Best Local Similarity 40.0%; Pred. No. 3.3;  
 Matches 16; Conservative 8; Mismatches 10; Indels 6; Gaps 1;

Qy 3 ARAVFLASQQLQARLMKKEES----PVYWSWRLEPEDG 36  
 Db 279 SRAMLLRFLAMEKGRRHMEYCPCPYLVYPVVAFRLEPKDG 318

RESULT 14  
IS-09-784-810A-2  
Sequence 2, Application US/09784810A  
Patent No. US802008220A1  
GENERAL INFORMATION:  
APPLICANT: RASTELLI, LUCA  
TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING  
TITLE OF INVENTION: SAME  
FILE REFERENCE: 10716-08  
CURRENT APPLICATION NUMBER: US/09/784,810A  
CURRENT FILING DATE: 2001-02-14  
PRIOR APPLICATION NUMBER: 60/182,360  
PRIOR FILING DATE: 2000-02-14  
PRIOR APPLICATION NUMBER: 60/191,261  
PRIOR FILING DATE: 2000-03-22  
NUMBER OF SEQ ID NOS: 29  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 384  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-784-810A-2  
Query Match Similarity 29.1%; Score 62; DB 9; Length 384;  
Best Local Similarity 40.0%; Pred. No. 3 4;  
Matches 16; Conservative 8; Mismatches 10; Indels 6;  
Gaps

3 AEAFLAQLQARLMEKES-----PYSWRLEPEDG 36  
   :|: | :| :| :| :| :| :| :| :| :| :| :| :|  
 295 SRAMILRLFLAMEKGHRMEYSCPVLVYVPAFRLEPKDG 334

Search completed: July 7, 2004, 17:14:58  
Search time: 48 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: July 7, 2004, 17:05:13 ; Search time 20 Seconds

(without alignments)  
206.812 Million cell updates/sec

Title: US-10-081-108-2

Perfect score: 213

Sequence: 1 MARRAVFIALLSAGOLLQARLM.....SPVVSRLPEPDGTALCFIF 43

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

searched:

283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR78:\*

1: Pir1:\*

2: Pir2:\*

3: Pir3:\*

4: Pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match	Length	ID	Description
1	55.5	26.1	359	2	JC7280	cytokine receptor-transposase - Esch
2	55.5	25.8	154	2	F69520	probable excrinucle
3	54.5	25.6	994	1	TOBCTS	hypothetical prote
4	54	25.4	295	2	G90934	conserved hypothetical prote
5	54	25.4	295	2	E64933	conserved hypothetical prote
6	54	25.4	295	2	C85783	conserved hypothetical prote
7	54	25.4	302	2	AD0709	conserved hypothetical prote
8	54	25.4	552	2	T00459	conserved hypothetical prote
9	54	25.4	552	2	AC0609	conserved hypothetical prote
10	54	25.4	699	2	A96802	unknown protein [i]
11	53.1	25.1	155	2	E97591	structure homolog
12	53	24.9	552	2	D64826	Ybd protein - Esc
13	53	24.9	552	2	B80749	hypothetical prote
14	53	24.9	552	2	F85599	hypothetical prote
15	52.5	24.6	243	2	T04489	protein kinase hom
16	52	24.4	346	2	T51377	hypothetical prote
17	52	24.4	445	2	T05639	hypothetical prote
18	52	24.4	945	2	A64714	helicase - Helicob
19	51	23.9	254	2	T09009	cdd3-like protein
20	51	23.9	284	2	G70732	probable thiosulfa
21	51	23.9	332	2	A99738	probable membrane
22	51	23.9	332	2	B85588	hypothetical prote
23	51	23.9	332	2	C64846	glycerol kinase (E)
24	51	23.9	524	1	S36175	ATP-stimulated glu
25	51	23.9	524	2	JN0606	probable ATP-depen
26	51	23.9	946	2	A71805	fibrinogen-binding
27	50.5	23.7	862	2	AC1214	hypothetical prote
28	50	23.5	58	2	S13979	gpl protein - Myco
29	50	23.5	270	2	A72800	

## RESULT 1

JC7280

cytokine receptor-like molecule-2 - mouse

C;Species: Mus musculus (house mouse)

C;Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 18-Aug-2000

C;Accession: JC7280

R;Hiroyama, T.; Iwama, A.; Morita, Y.; Nakamura, Y.; Shibuya, A.; Nakauchi, H.

Biochem Biophys Res Commun. 272, 224-229, 2000

A;Title: Molecular cloning and characterization of CRLM-2, a novel type I cytokine rece

A;Reference number: JC7280

A;Contents: Embryo

A;Accession: JC7280

A;Molecule type: mRNA

A;Residues: 1-359 &lt;HR&gt;

A;Cross references: DDBJ:AB039945

C;Generics:

A;Gene: crlm2

C;Keywords: cytokine; embryo; receptor; signal transduction; transmembrane protein

Query Match % Score 55.5; DB 2; Length 359;

Best Local Similarity 44.0%; Pred. No. 9.6%; Mismatches 4; Indels 3; Gaps 1;

Matches 11; Conservative 7; Mismatches 4; Indels 3; Gaps 1;

Query Match % Score 55.5; DB 2; Length 359;

Best Local Similarity 44.0%; Pred. No. 9.6%; Mismatches 4; Indels 3; Gaps 1;

Matches 11; Conservative 7; Mismatches 4; Indels 3; Gaps 1;

Query Match % Score 55.5; DB 2; Length 359;

Best Local Similarity 44.0%; Pred. No. 9.6%; Mismatches 4; Indels 3; Gaps 1;

Matches 11; Conservative 7; Mismatches 4; Indels 3; Gaps 1;

Query Match % Score 55.5; DB 2; Length 359;

Best Local Similarity 44.0%; Pred. No. 9.6%; Mismatches 4; Indels 3; Gaps 1;

Matches 11; Conservative 7; Mismatches 4; Indels 3; Gaps 1;

Query Match % Score 55.5; DB 2; Length 359;

Best Local Similarity 44.0%; Pred. No. 9.6%; Mismatches 4; Indels 3; Gaps 1;

Matches 11; Conservative 7; Mismatches 4; Indels 3; Gaps 1;

Query Match % Score 55.5; DB 2; Length 359;

Best Local Similarity 44.0%; Pred. No. 9.6%; Mismatches 4; Indels 3; Gaps 1;

Matches 11; Conservative 7; Mismatches 4; Indels 3; Gaps 1;





Science 277, 1453-1462, 1997  
*A;Title:* The complete genome sequence of *Escherichia coli* K-12.  
*A;Reference:* A64720; MUID:7426617; PMID:9278503  
*A;Status:* nucleic acid sequence not shown; translation not shown  
*A;Molecule type:* DNA  
*A;Residues:* 1-552 <BLAT>  
*A;Cross-references:* GB:AE000189; GB:U0096; NID:91787097; PIDN: AAC73963.1; PID:91787102;  
*A;Experimental source:* strain K-12, substrate MG1655  
*C;Genetics:*  
*A;Gene:* ybdJ  
*C;Keywords:* ATP; nucleotide binding; P-loop; transmembrane protein  
*F:29-36/Region:* nucleotide-binding motif A (P-loop)  
*F:264-280/Domain:* transmembrane #status predicted <TM>

Query Match Score 24.9%; Best Local Similarity 34.9%; Pred. No. 33; Matches 15; Conservative 5; Mismatches 11; Indels 12; Gaps 2;

Qy 11 SAQLIQ----ARLMKERSPVVSWRLSP----EDGTAFCF 41  
Db 324 SGELSLSLTPVEHVCRVLRSSRVAAWRLGPSGLSTEDSRRI SF 366

RESULT 15

T04489 protein kinase homolog F8F16.40 - *Arabidopsis thaliana*  
*C;Species:* *Arabidopsis thaliana* (mouse-ear cress)  
*C;Date:* 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 30-Apr-1999  
*C;Accession:* T04489  
*R;Bevan, M.; Brandt, P.; Dose, S.; Jarke, D.; Scharfe, M.; Schon, O.; Hoheisel, J.; Meier, A;Reference number:* 215375  
*A;Accession:* T04489  
*A;Molecule type:* DNA  
*A;Residues:* 1-243 <EBV>  
*A;Cross-references:* EMBL:AL021633  
*A;Experimental source:* cultivar Columbia; BAC clone F8F16  
*C;Genetics:*  
*A;Map Position:* 4  
*A;Note:* F8F16.40

B90749 hypothetical protein Ec0962 [Imported] - *Escherichia coli* (strain O157:H7, substrate R1)  
*C;Species:* *Escherichia coli*  
*C;Accession:* 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 18-Jul-2001  
*C;Accession:* B90749  
*R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.*  
*DNA Res.* 8, 11-22, 2001  
*A;Title:* Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and geno

*A;Accession:* A99629; MUID:21156231; PMID:11258796  
*A;Status:* preliminary  
*A;Molecule type:* DNA  
*A;Residues:* 1-552 <HAY>  
*A;Cross-references:* GB:BA000007; PIDN:BAB34385.1; PID:913360421; GSPDB:GN00154  
*A;Experimental source:* strain O157:H7, substrate RIMD 0509952  
*C;Genetics:*  
*A;Gene:* ECs0962

RESULT 13

B90749 hypothetical protein Ec0962 [Imported] - *Escherichia coli* (strain O157:H7, substrate R1)  
*C;Species:* *Escherichia coli*  
*C;Accession:* 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 18-Jul-2001  
*C;Accession:* B90749  
*R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.*  
*DNA Res.* 8, 11-22, 2001  
*A;Title:* Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and geno

*A;Accession:* B90749  
*A;Status:* preliminary  
*A;Molecule type:* DNA  
*A;Residues:* 1-552 <HAY>  
*A;Cross-references:* GB:BA000007; PIDN:BAB34385.1; PID:913360421; GSPDB:GN00154  
*A;Experimental source:* strain O157:H7, substrate RIMD 0509952  
*C;Genetics:*  
*A;Gene:* ECs0962

Query Match Score 24.9%; Best Local Similarity 34.9%; Pred. No. 33; Matches 15; Conservative 5; Mismatches 11; Indels 12; Gaps 2;

Qy 11 SAQLIQ----ARLMKERSPVVSWRLSP----EDGTAFCF 41  
Db 324 SGELSLSLTPVEHVCRVLRSSRVAAWRLGPSGLSTEDSRRI SF 366

Search completed: July 7, 2004, 17:09:33  
Job time : 22 secs

RESULT 14

F85599 hypothetical protein ybdJ [Imported] - *Escherichia coli* (strain O157:H7, substrate EDL93)  
*C;Species:* *Escherichia coli*  
*C;Accession:* 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
*C;Accession:* F85599  
*R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, L.; Grotbeck, B.J.; Davis, N.W.; Lin, A.; Dimaranta, E.; Potamousis, K.; Apodaca, Nature 401, 529-533, 2001*  
*A;Title:* Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.  
*A;Accession:* A85480; MUID:21074935; PMID:11206551  
*A;Status:* preliminary  
*A;Molecule type:* DNA  
*A;Residues:* 1-552 <STO>  
*A;Cross-references:* GB:AE005174; NID:912511906; PIDN:AAG55258.1; GSPDB:GN00145; UWGP:Z11  
*A;Experimental source:* strain O157:H7, substrate EDL93  
*C;Genetics:*  
*A;Gene:* ybdJ

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2004, 16:58:33 ; Search time 12 Seconds  
(without alignments)  
186.385 Million cell updates/sec

Title: US-10-081-108-2  
Perfect score: 213  
Sequence: 1 MAARAYFLALSAQILQARLM.....SPVVSWRLEPEDGTLACPFIF 43

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%Listing first 45 summaries  
Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	213	100.0	43	1	BGE1_HUMAN	Q13072 homo sapien	013072 homo sapien
2	206	96.7	43	1	BGE5_HUMAN	Q86527 homo sapien	Q86527 homo sapien
3	177	83.1	109	1	BGE2_HUMAN	Q86730 homo sapien	Q86730 homo sapien
4	177	83.1	109	1	BGE3_HUMAN	Q86729 homo sapien	Q86729 homo sapien
5	173	81.2	39	1	BGE4_HUMAN	Q86528 homo sapien	Q86528 homo sapien
6	62	29.1	384	1	SPH1_HUMAN	Q9yai1 homo sapien	Q9yai1 homo sapien
7	55.5	26.1	359	1	CRL2_MOUSE	Q8cji9 mus musculu	Q8cji9 mus musculu
8	55.5	25.8	154	1	Y16_1_ARCFU	Q28116 archaeoglob	Q28116 archaeoglob
9	54	25.6	994	1	TNP5_ECOLI	P08514 escherichia	P08514 escherichia
10	54	25.4	293	1	CHO_SALTY	Q8z595 salmonella	Q8z595 salmonella
11	54	25.4	293	1	CHO_ECO57	Q8zpu6 salmonella	Q8zpu6 salmonella
12	54	25.4	295	1	CHO_ECOL6	P59x27 escherichia	P59x27 escherichia
13	54	25.4	295	1	CHO_ECOL6	P59439 escherichia	P59439 escherichia
14	54	25.4	295	1	CHO_ECOLI	P76213 escherichia	P76213 escherichia
15	54	25.4	295	1	CHO_SHAF1	P59x31 shigella f1	P59x31 shigella f1
16	54	25.4	1080	1	HDA4_CHICK	P283038 shigella f1	P283038 shigella f1
17	54	25.4	1084	1	HDA4_HUMAN	P256524 homo sapien	P256524 homo sapien
18	53	24.9	552	1	YBJD_ECOLI	P75828 escherichia	P75828 escherichia
19	52.5	24.6	170	1	YC66_MESVI	Q9mug8 mesotigma	Q9mug8 mesotigma
20	52	24.4	332	1	YBHG_ECOL6	Q8f616 escherichia	Q8f616 escherichia
21	51	23.9	284	1	THT3_MYCTU	Q959570 mycobacteri	Q959570 mycobacteri
22	51	23.9	332	1	YBHG_ECOL7	Q8xty9 escherichia	Q8xty9 escherichia
23	51	23.9	332	1	YBHG_ECOLI	P75777 escherichia	P75777 escherichia
24	51	23.9	332	1	YBHG_SHFL	Q833616 shigella f1	Q833616 shigella f1
25	51	23.9	457	1	CD4_SALIC	Q29037 sainiri sci	Q29037 sainiri sci
26	51	23.9	524	1	GLPK_HUMAN	P32189 homo sapien	P32189 homo sapien
27	51	23.9	524	1	GLPK_MOUSE	Q64516 mus musculu	Q64516 mus musculu
28	51	23.9	524	1	GLPK_RAT	Q63060 rattus norv	Q63060 rattus norv
29	51	23.9	1113	1	HDA5_MOUSE	Q32276 mus musculu	Q32276 mus musculu
30	50	23.5	270	1	VGP1_BMD2	Q64197 mycobacteri	Q64197 mycobacteri
31	50	23.5	708	1	TPAB_HUMAN	Q9ygi0 homo sapien	Q9ygi0 homo sapien
32	50	23.5	1122	1	HDAS_HUMAN	Q9uq16 homo sapien	Q9uq16 homo sapien
33	49.5	23.2	699	1	MIB1_HUMAN	Q9ukm7 homo sapien	Q9ukm7 homo sapien

## ALIGNMENTS

34	49.5	23.2	821	1	FGR2_HUMAN	P21802 homo sapien				
35	49	23.0	351	1	FMSY_MYCHO	O32861 mycoplasma				
36	49	23.0	458	1	CD4_HUMAN	P01730 homo sapien				
37	49	23.0	458	1	CD4_PANTR	P16004 pan troglod				
					Q94738	strongly loco				
					38	23.0	886	1	HS97_SRPN	Q06065 strongly loco
					39	23.0	889	1	HS97_SRPN	P05703 escherichia
					40	22.8	101	1	CCDB_ECOLI	Q52042 escherichia
					41	22.8	126	1	CCB3_ECOLI	Q52043 escherichia
					42	22.8	126	1	YIEL_ECOLI	P31471 escherichia
					43	22.8	400	1	PPA1_ASPPNG	P20584 aspergillus
					44	22.8	436	1	OSTE_SCCHPO	O14238 schizosacch
					45	22.5	122	1		

RESULT 1  
BGE1\_HUMAN  
ID BGE1\_HUMAN  
AC Q13072\_HUMAN  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE B melanoma antigen 1 precursor (B melanoma antigen) (Antigen MZ2-BA).  
GN BAGE OR BGE1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Buteraria; Primates; Catarrhini; Hominidae; Homo.  
NCBI\_TAXID=1606;  
RN [1]  
RP RP SEQUENCE FROM N.A. (ISOFORM BAGE1A).  
RC TISSUE-Melanoma;  
RX MEDLINE=95202592; PubMed=7895173;  
RA Boel P., Waldmann C., Sensi M.L., Brasseur R., Renaud J.-C., De Sario A., Coulie P., Bonn T., van der Bruggen P.; "BAGE: a new gene encoding an antigen recognized on human melanomas by cytolytic T lymphocytes." Eur. J. Hum. Genet. 10:833-840 (2002).  
RT RT ALTERNATIVE SPlicing, Named isoforms=1;  
RN [2]  
RP ALTERNATIVE SPlicing.  
RX MEDLINE=22349465; PubMed=12461691;  
RA Ruaubat P., Brun M.-E., Boyle S., Roizes G., De Sario A., "New BAGE (B melanoma antigen) genes mapping to the juxtacentromeric regions of human chromosomes 13 and 21 have a cancer/testis expression profile." Eur. J. Hum. Genet. 10:833-840 (2002).  
RN [3]  
RP FUNCTION:Unknown. Antigen recognized on a melanoma by autologous cytolytic T lymphocytes.  
CC CC -1 ALTERNATIVE PRODUCTS:  
RN Comment=At least 6 different mRNAs, BAGE1a, BAGE1b, BAGE1c, BAGE1d, BAGE1e and BAGE1f, are produced by alternative splicing. They differ in the 3' region, but give rise to the same protein sequence;  
CC CC IsoID=C13072-1; Sequence=Displayed;  
RN [4]  
RP TISSUE SPECIFICITY: Not expressed in normal tissues, except in testis. Expressed with significant proportion in melanomas, but also in tumors of various histological origins, such as bladder carcinomas, head and neck squamous cell carcinomas, lung and breast carcinomas. Not expressed in renal, colorectal and prostatic carcinomas, leukemias and lymphomas. More frequently expressed in metastatic melanomas than in primary melanomas.  
CC CC -1 MISCELLANEOUS: The ancestral BAGE gene was generated by juxtapacentromeric reshuffling of the ML3 Gene. The BAGE family was expanded by juxtapacentromeric movement and/or acrocentric exchanges. BAGE family is composed of expressed genes that map to the juxtapacentromeric regions of chromosomes 13 and 21 and of unexpressed gene fragments that scattered in the juxtapacentromeric regions of several chromosomes, including chromosomes 9, 13, 18 and 21.

CC -!- SIMILARITY: Belongs to the BAGE family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC use by non-profit institutions. There are no restrictions on its

CC modifiability and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

DR EMBL: U19180; AAC50123.1; -

DR EMBL: AAP0743.1; -

DR EMBL: AF521550; AAQ2670.1; -

DR EMBL: AF521551; AAC012671.1; -

DR EMBL: AF521552; AAC012672.1; -

DR EMBL: AF521553; AAC012673.1; -

DR EMBL: AF521554; AAC012674.1; -

DR Genew: HGNC: 942; BAGE.

DR NM\_605171; -.

KW Antigen; Alternative splicing; Multigene family; Signal.

FT SIGNAL 17 POTENTIAL.

FT CHAIN 18 43 AA; 4810 MW; 36F3BCB012FLBB CRC64;

SQ SEQUENCE 43 AA; 4810 MW; C6F3ACE4012FLA6 CRC64;

Query Match 100.0%; Score: 213; DB 1; Length: 43;

Best Local Similarity 100.0%; Pred. No. 1..e-22;

Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAARAVFLAQLAQQLQARLNKEESVSVRLPEPDGTALCFIF 43

Db 1 MAARAVFLAQLAQQLQARLNKEESVSVRLPEPDGTALCFIF 43

RESULT 3

BGE2\_HUMAN STANDARD; PRT; 109 AA.

ID BGE2\_HUMAN Q86Y30; AC 10-0CT-2003 (Rel. 42, Created)

ID BGE2\_HUMAN Q86Y30; DT 10-OCT-2003 (Rel. 42, Last sequence update)

AC Q86Y30; DT 10-OCT-2003 (Rel. 42, Last annotation update)

AC Q86Y30; DE B\_melanoma\_antigen\_2 precursor.

AC Q86Y30; GN BAGE2.

AC Q86Y30; OS Homo\_sapiens\_(Human).

AC Q86Y30; OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AC Q86Y30; OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AC Q86Y30; OC NCBITaxID=9606;

RN [1] RN SEQUENCE FROM N.A.

RC TISSUE=Testis; RX MEDLINE=2234945; PubMed=12451691;

RN [1] RA Ruauit M., Van Der Bruggen P., Brun M.-E., Boyle S., Roizes G.,

RC SEQUENCE FROM N.A. RA Sarro A.D.; RT New BAGE (B\_melanoma antigen) genes mapping to the juxtapacentromeric

RX TISSUE=Testis; RT regions of human chromosomes 13 and 21 have a cancer/testis expression

RN [1] RA Ruauit M., Ventura M., Galtier N., Brun M.-E., Archidiacono N., profile."; RT profile.";

RC SEQUENCE FROM N.A. RL Eur. J. Hum. Genet. 10:833-840(2002).

RX TISSUE=Testis; RX MEDLINE=22563396; PubMed=12676563;

RN [1] RA Ruauit M., Ventura M., Galtier N., Brun M.-E., Archidiacono N.,

RC SEQUENCE FROM N.A. RA Roizes G., De Sarro A.; RT "BAGE genes generated by juxtapacentromeric reshuffling in the hominidae

RA Ruauit M., Ventura M., Galtier N., Brun M.-E., Archidiacono N., lineage are under selective pressure.";

RA Sarro A.D.; RT Genomics 81:391-399(2003).

RT "New BAGE (B\_melanoma antigen) genes mapping to the juxtapacentromeric regions of human chromosomes 13 and 21 have a cancer/testis expression profile."; RT

RL Genomics 81:391-399(2003).

RN [2] RA Ruauit M., Ventura M., Galtier N., Brun M.-E., Archidiacono N.,

RC SEQUENCE FROM N.A. RA Roizes G., De Sarro A.; RT "BAGE genes generated by juxtapacentromeric reshuffling in the hominidae

CC -!- FUNCTION: Unknown. Candidate gene encoding tumor antigens. CC -!- TISSUE SPECIFICITY: Not expressed in normal tissues except in CC testis. Expressed in 22% of melanomas, in bladder and lung carcinomas.

CC -!- MISCELLANEOUS: The ancestral BAGE gene was generated by CC juxtapacentromeric reshuffling of the MLL3 gene. The BAGE family was CC composed of expressed genes that map to the juxtapacentromeric regions of chromosomes 13 and 21 and of CC unexpressed gene fragments that map to the juxtapacentromeric regions of chromosomes 13 and 21. CC

- - SIMILARITY: Belongs to the BAGE family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license/agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to license@isb-sib.ch).

CC DR EMBL; AF218570; AAL55648 ; - .

CC DR Genew; HGNC:15723; BAGE2 ; - .

CC KW Antigen; Multigene family; Signal.

CC SIGNAL 1 17 POTENTIAL.

CC FT CHAIN 18 109 B MELANOMA ANTIGEN 2 .

CC SEQUENCE 109 AA; 12114 MW; 51797BAA2C135FF CRC64 ;

CC SQ

Query Match Score 177; DB 1; Length 109;

Best Local Similarity 94.9%; Prod. No. 2\_7e-17;

Matches 37; Conservative 0; Missmatches 2; Indels 0; Gaps 0;

Qy 1 M A R A V A F I L A S Q A L Q I Q A R I M K E S P V V S W R L E P D G T A L 39

Db 1 M A J A G V F V F L A S Q A L Q I Q A R I M K E S P V V S W R L E P D G T A L 39

RESULT 4

AC	LOCUS	DEFINITION	VERSION	COMMENT
BGE3_HUMAN	ID BGE3_HUMAN	STANDARD	ERT;	109 AA.
Q86Y29	DT 10-OCT-2003 (Rel. 42, Created)			
	DT 10-OCT-2003 (Rel. 42, Last sequence update)			
	DT 10-OCT-2003 (Rel. 42, Last annotation update)			
	DE melanoma antigen 3 precursor.			
	GN BAGE3.			
	OS Homo sapiens (Human).			
	OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
	RN [1] _TAXID=9606;			
	RP SEQUENCE FROM N.A.			
	RX MEDLINE:2239465; PubMed=12461691;			
	RA Rueault M., Van Der Bruggen P., Brun M.-E., Boivie S., Roizes G.,			
	RA Sario A.D.,			
	RT "New BAGE (B melanoma antigen) genes mapping to the juxtacentromeric regions of human chromosomes 13 and 21 have a cancer/testis expression profile.";			
	RT Bur. J. Hum. Genet. 10:833-840 (2002).			
	RL [2]			
	RP SEQUENCE FROM N.A.			
	RX MEDLINE:22563396; PubMed=12576563;			
	RA Rueault M., Ventura M., Galtier N., Brun M.-E., Archidiacono N.,			
	RA Roiz G., De Sario A.,			
	RT "BAGE genes generated by juxtacentromeric reshuffling in the hominidae lineage are under selective pressure.";			
	RT Genomics 81:391-399 (2003).			
	RU [3]			
	CC -FUNCTION: Unknown. Candidate gene encoding tumor antigens.			
	CC -TISSUE SPECIFICITY: Not expressed in normal tissues except in			
	CC testis. Expressed in melanoma, bladder and lung carcinomas.			
	CC -MISCELLANEOUS: The ancestral BAGE gene was generated by juxtacentromeric reshuffling of the MLL3 gene. The BAGE family was expanded by juxtacentromeric movement and/or acrocentric exchanges. BAGE family is composed of expressed genes that map to the juxtacentromeric regions of chromosomes 13 and 21 and of unexpressed gene fragments that scattered in the juxtacentromeric regions of several chromosomes, including chromosomes 9, 13, 18 and 21.			
	CC -SIMILARITY: Belongs to the BAGE family.			

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

---

DR EMBL; AF319514; AAO32634; 1; -.

DR GenBank; HGNC:15728; BAGE3;

KW Antigen; Multigen Family; Signal.

KW POTENTIAL.

FT SIGNAL\_1

FT CHAIN\_B MELANOMA ANTIGEN 3.

FT SEQUENCE\_AA: 109 MW: A477B7A8FC3A2F4E CRC64;

FT SEQUENCE\_109 AA: 12112

Query Match 83.1%; Score 177; DB 1; Length 109;

Best Local Similarity 94.9%; Pred. No. 2.7e-17;

Matches 37; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 1 MAARAVFLAISAQQLQARLIMKEESPVYWSRLEPEDGTAI 39

Db 1 MAAGVYFLAISAQQLQARLIMKEESPVYWSRLEPEDGTAI 39

---

RESULT 5

	BGE4_HUMAN	STANDARD	PRT;	39 AA.
ID	BGE4_HUMAN			
QB6YY8;				
DT	10-OCT-2003	(Rel. 42, Created)		
DT	10-OCT-2003	(Rel. 42, Last sequence update)		
DT	10-OCT-2003	(Rel. 42, Last annotation update)		
DE	B melanoma antigen 4 precursor.			
GN	BAGE4.			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID:9606;			
RN	[1] —			
RP	SEQUENCE FROM N.A.			
RC	SEQUENCE=Testis;			
RA	Ruault M., Rocchi M., Boyle S., Roizes G., van der Bruggen P., De Sario A.;			
RA	"BAGE a family of centromeric genes coding for tumor antigens.";			
RT	Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.			
RL	- - FUNCTION: Unknown. Candidate gene encoding tumor antigens.			
CC	- - TISSUE SPECIFICITY: Not expressed in normal tissues except in testis. Expressed in melanoma, bladder and lung carcinomas.			
CC	- - MISCELLANEOUS: The ancestral BAGE gene was generated by juxtacentromeric reshuffling of the MLL3 gene. The BAGE family was expanded by juxtacentromeric movement and/or acrocentric exchanges. BAGE family is composed of expressed genes that map to the juxtacentromeric regions of chromosomes 13 and 21 and of unexpressed gene fragments that scattered in the juxtapacentromeric regions of several chromosomes, including chromosomes 9, 13, 18 and 21.			
CC	- - SIMILARITY: Belongs to the BAGE family.			

---

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

---

DR EMBL; AF319515; AAO32635; 1; -.

DR GenBank; HGNC:15730; BAGE4

Db	MAAGAFLAASQQLQARLMKEESPVVSWEPEPDGTAL	39	CC entities requires a license agreement. (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).
RESULT 6			
SPPH1 HUMAN	STANDARD; PRT;	384 AA.	
ID SPHK1	Q9NYA2; Q9HD92; Q9NY70; Q9NYL3;		
AC	Q9NYA2; Q9HD92; Q9NY70; Q9NYL3;		
DT 28-FEB-2003	(Rel. 41; Created)		
DT 28-FEB-2003	(Rel. 41; Last sequence update)		
DT 10-OCT-2003	(Rel. 42; Last annotation update)		
DB Sphingosine kinase 1	(EC 2.7.1.-) (SK 1)		
GN SPHK1 OR SPHK OR SPK.			
OS Homo sapiens (Human).			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;			
OC Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.			
NCBI_TaxID=9606;			
RN [1]			
RP SEQUENCE FROM N.A.			
RX MEDLINE=20323213; PubMed=10863092;			
RA Melendez A.J., Carlos-Dias E., Gosink M., Allen J.M., Takacs L.;			
RT "Human sphingosine kinase: molecular cloning, functional characterization and tissue distribution."			
RL Gene 25(1):19-26(2000).			
RN [2]			
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.			
RX MEDLINE=20263793; PubMed=10802064;			
RA Nava V.E., Lacuna E., Poulet S., Liu H., Sugiyama M., Kono K.,			
RA Milstein S., Kohama T., Spiegel S.,			
RT "Functional characterization of human sphingosine kinase-1."			
RL FEBS Lett. 473(81-84)(2000).			
RN [3]			
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.			
RX MEDLINE=20407120; PubMed=10937957;			
RA Pitson S.M., D'Andrea R.J., Vandelleur L., Moretti P.A.B., Xia P.,			
RA Gamble J.R., Vadas M.A., Wattenberg B.W.;			
RT "Human sphingosine kinase: purification, molecular cloning and characterization of the native and recombinant enzymes."			
RL Biochem. J. 350(429-441)(2000).			
RN [4]			
RP SEQUENCE FROM N.A.			
RA Van Veldhoven P.P., Gijssbers S.; Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.			
RN [5]			
RP SEQUENCE FROM N.A.			
RC TISSUE=Mammary gland, and Ovary;			
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,			
RA Nishikawa T., Nagai K., Sugai S., Shotoraki A., Sudo H., Kondo H., Sugawara M.,			
RA Watanabe M., Hosoi T., Kaku Y., Koaira H., Kondo H., Sugawara M.,			
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,			
RA Yamamoto S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,			
RA Ninomiya K., Iwayanagi T., Nakamura Y., Nagahari K., Masuho Y.,			
RT "NEDO human cDNA sequencing project."			
RL Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.			
CC - FUNCTION: Catalyzes the phosphorylation of sphingosine to form sphingosine 1-phosphate (SPP), a lipid mediator with both intra- and extracellular functions. Also acts on D-erythro-sphingosine and to a lesser extent sphinganine, but not other lipids, such as D,L-threo-dihydroxyphingosine, N,N-dimethylsphingosine, diacylglycerol, ceramide, or phosphatidylinositol.			
CC - CATALYTIC ACTIVITY: Sphingosine + ATP = sphingosine 1-phosphate + ADP.			
CC - SUBUNIT: Binds to calmodulin.			
CC - TISSUE-SPECIFICITY: Widely expressed with highest levels in adult liver, kidney, heart and skeletal muscle.			
CC - SIMILARITY: Contains 1 DAGKC domain.			
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial			
CC SSEQUENCE FROM N.A. (ISCBFORMS 1 AND 2).			
RC TISSUE=Embryo;			
RC MEDLINE=20329312; PubMed=10872831;			
RA Hirovama T., Iwama A., Morita Y., Nakamura Y., Shibuya A., Nakachi H.;			
RA "Molecular cloning and characterization of CRLM-2, a novel type I			
NCBI_TaxID=10000;			
RN [1] -			

cytokine receptor preferentially expressed in hematopoietic cells.";  
RJ Biochem. Biophys. Res. Commun. 272:224-229(2000).  
[2] SEQUENCE FROM N.A. (ISOFORM 1), AND ALTERNATIVE SPlicing.  
RC TISSUE=lymphocytes;  
RX MEDLINE=20197866; PubMed=110733486; RFL  
RA Fujio K., Nobuka T., Kojima T., Kawashima T., Yahata T., Matsuura T.,  
RA Copeland N.G., Gilbert D.J., Jenkins N.A., Yamamoto K.,  
RA Kitamura T.;  
RT "Molecular cloning of a novel type I cytokine receptor similar to the common gamma chain.";  
RL Blood 95:2204-2211(2000).  
[3] RPP SEQUENCE FROM N.A. (ISOFORM 1).  
RC STRAIN=C57BL/6J; TISSUE=lymphocytes;  
RX MEDLINE=20432254; PubMed=110974032;  
RA Park L.S., Martin J., Garka K., Gliniak B., Di Santo J.P., Muller W.,  
RA Ziegler S.F., Morrissey P.J., Paxton R., Sims J.B.;  
RA "Cloning of the murine thymic stromal lymphopoietin (TSLP) receptor."  
RT Formation Of a functional heteromeric complex requires interleukin-7 receptor";  
RL J. Exp. Med. 192:659-670(2000).  
RN [4] RPP SEQUENCE FROM N.A. (ISOFORM 1).  
RC STRAIN=FVB/N;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Straubberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,  
RA Klausner R.D., Collins F.S., Wagner L., Shemesh C.M., Schuler G.D.,  
RA Altzinger S.F., Zeeberg B.R., Buston K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Scarce M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Yoshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEvily P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmitz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.B.,  
RA Schnech A., Schein J.E., Jones S.J.M., Marra M.A.;  
RA "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";  
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [6] RPP SEQUENCE OF 234-359 FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=embryonic stem cells;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Aizawa K., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Saito T., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamakada I.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Saito R.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
RA Schirali L.M., Staibli F., Suzuki R., Tomita M., Wigner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujit M., Gariboldi M.,  
RA Gustincovich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Monbaerts P.,  
RA Nozono P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyoda K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
RA Wynnshaw-Borod A., Yoshida K., Hasegawa Y., Kawaji H.,  
RA Hayashizaki Y.;  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
CC -!- FUNCTION: Receptor for thymic stromal lymphopoietin (TSLP). Forms a functional complex with TSLP and IL7R which is capable of stimulating cell proliferation through activation of STAT3 and STAT5. Also activates JAK2. Implicated in the development of the hematopoietic system.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (isoform 1 and isoform 3). Secreted (isoform 2).  
CC -!- ALTERNATIVE PRODUCTS:  
CC -!- Event=Alternative splicing; Named isoforms=3;  
Comment=Additional isoforms seem to exist;  
Name=1; IsoID=Q8CII9-1; Sequence=Q8CII9-1; SequenceDisplayed:  
CC Name=2; Synonyms=Soluble CRM192;  
IsoID=Q8CII9-2; Sequence=VSP\_008788; VSP\_008789;  
Name=3; IsoID=Q8CII9-3; Sequence=VSP\_008790;  
CC Note=No experimental confirmation available;  
CC -!- TISSUE SPECIFICITY: High level of expression in liver, lung and testis. Also expressed in heart, brain, spleen, thymus and bone marrow. Highly expressed in progenitors and myeloid cells. Isoform 2 is expressed in primary hematopoietic cells.  
CC -!- INDUCTION: Upregulated in the myeloid 32D cell line by granulocyte colony stimulating factor (G-CSF).  
CC -!- DOMAIN: The WSXWS motif appears to be necessary for proper protein folding and thereby efficient intracellular transport and cell-surface receptor binding.  
CC -!- SIMILARITY: Belongs to the type I cytokine family of receptors.  
CC -!- SIMILARITY: Contains 1 fibronectin type III domain.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce> or send an email to license@isb-sib.ch).  
CC DR AB039945; BAA92684.1; DR PROSITE; PS03355; HEMATOPOIETIC\_SF1; FALSE\_NEG.  
CC DR AB031333; BAA92359.1; DR PROSITE; PS01355; TRANSMEMbrane; Glycoprotein; Alternative splicing.  
CC DR AB232936; AAF81676.1; DR SIGNAL; Signal; Transmembrane; Glycoprotein; Alternative splicing.  
CC DR AF01953; AAF82189.1; DR CYTOKINE RECEPTOR-LIKE FACTOR\_2.  
CC DR BC023738; AAH23788.1; DR SIGNAL; Signal; Transmembrane; Glycoprotein; Alternative splicing.  
CC DR AK010291; BAB26827.1; DR CYTOSOLIC (POTENTIAL).  
CC DR MG1:188506; TS1P; DR FIBRONECTIN\_TYPE-III.  
CC DR InterPro; IPR03961; FN\_III.  
CC DR Pfam; PF00041; fn3; 1.  
CC DR SMART; SM0060; FN2; 1.  
CC DR PROSITE; PS01355; HEMATOPOIETIC\_SF1; FALSE\_NEG.  
KW Receiver; Signal; Transmembrane; Glycoprotein; Alternative splicing.  
FT SIGNAL; Signal; Transmembrane; Glycoprotein; Alternative splicing.  
FT CHAIN 19 359 CYTOKINE RECEPTOR-LIKE FACTOR\_2.  
FT DOMAIN 20 232 EXTRACELLULAR (POTENTIAL).  
FT DOMAIN 233 253 POTENTIAL.  
FT DOMAIN 254 359 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 117 208 FIBRONECTIN\_TYPE-III.  
FT SITE 201 205 WSXWS motif.  
FT SITE 262 269 BOX\_1.  
FT DISULFHYD 68 82 BY SIMILARITY.  
FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 122 122 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT VARSPLIC 217 217 A > G (In isoform 2).  
FT VARSPLIC /FTD-VSP 008788. Missing (In isoform 2).  
FT VARSPLIC 218 359 /FTD-VSP 008789. A > AGDPAAHNP (In isoform 3).  
FT VARSPLIC 217 217

FT	CONFLICT	15	15	/FTID=wsp_008790.	
FT	CONFLICT	51	51	A -> T (IN REF. 1).	
FT	CONFLICT	87	87	G -> S (IN REF. 4 AND 5).	
FT	CONFLICT	179	179	A -> G (IN REF. 4).	
FT	CONFLICT	309	309	A -> V (IN REF. 1 AND 3).	
SQ	SEQUENCE	359 AA;	37761 MW;	F9C521C5B4AC9D CRC64;	
Query Match		26.1%	Score 55.5;	DB 1; Length 359;	
Best Local Similarity		44.0%	Pred. No. 4.6;		
Matches	11;	Conservative	7;	Mismatches	
OY		21	KEEPSVW---SWRLLEPDGDTALCFI 42		
Db		301	EEDDLILPKRKVEPBDGTSLCTV 325		
RESULT 8					
YI166_ARCFU	STANDARD;	PRT;	154 AA.		
AC	028116;				
DT	16-OCT-2001	(Rel. 40, Created)			
DT	16-OCT-2001	(Rel. 40, Last sequence update)			
DT	16-OCT-2001	(Rel. 40, Last annotation update)			
DE	Hypothetical Protein AF2166.				
GN	AF2166.				
CC	Archaeoglobus fulgidus.				
CC	Archaeoglobaceae; Archaeoglobi; Archaeoglobales;				
CC	Archaeoglobobacae; Archaeoglobob.				
CX	NCBI_TaxID=2234;				
RN	[1]				
SEQUENCE FROM N.A.					
RC	STRAIN=VC-16 / DSM 4304 / ATCC 49558;				
RX	MEDLINE=98049343; PubMed=9388475;				
RA	Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E., Richardson D.J., Dodson R.J., Kyrpides N.C., Peterson J.D., Ketchum K.A., Kerlavage A.R., Graham D.B., Hickey E.K., Peterson K.E., Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S., Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B., Peterson S., Deboy R.T., McNeil L.K., Badger J.H., Godek A., Zhou L., Overbeek R., Gocayne J.D., Weidman J.P., McDonald L., Utterback T., Cottrell M.D., Spriggs T., Ariyachai P., Sodow P.W., D'Andrea K.P., Bowman C., Fujii C., Sykes S.M., Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woose C.R., Venter J.C.;				
RA	"The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon <i>Archaeoglobus fulgidus</i> .";				
RT	Nature 390:364-370 (1997).				
RC	- - SUBCELLULAR LOCATION: Integral membrane protein (Potential).				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).				
DR	EMBL; AB000955; AAB89089.1; - .				
DR	TIGR; AF2166; - .				
KW	Hypothetical Protein; Transmembrane; Complete proteome.				
FT	TRANSMEM 15 37 POTENTIAL.				
FT	TRANSMEM 58 80 POTENTIAL.				
FT	TRANSMEM 95 116 POTENTIAL.				
FT	TRANSMEM 123 145 POTENTIAL.				
SEQUENCE	154 AA; 17266 MW; 1CE4.1261CD8FA468 CRC64;				
Query Match	25.8%	Score 55;	DB 1; Length 154;		
Best Local Similarity	34.2%	Pred. No. 2.2;			
Matches	13;	Conservative	10;	Mismatches	
OY	2 AARAVFLAQLQARKEESPVWRLEDGTAL 39				

OC Enterobacteriaceae; Salmonella.  
 OC [1] NCBI\_TaxID=603;  
 RN OC  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CT18;  
 XX MEDLINE=21534947; PubMed=11677608;  
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D.J., Wain J.,  
 RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,  
 RA Baker S., Brooks K., Chillingworth T., Connerton P.,  
 RA Cronin A., Davies P.M., Dowd L., White N., Farrar J.,  
 RA Feltwell T., Hamlin N., Haque A., Hien T.N., Holroyd S., Jigels K.,  
 RA Krogh A., Larsen T.S., Leather S., McEne S., O'Gaora P., Parry C.,  
 RA Quail M.A., Rutherford K., Simmonds M., Skeleton J., Stevens K.,  
 RA Whitehead S., Barrell B.G.; "Complete genome sequence of a multiple drug resistant *Salmonella* enterica serovar Typhi CT18.", Nature 413:848-852 (2001).  
 [2] RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Ty2 / ATCC 700931;  
 RX MEDLINE=21534367; PubMed=12644504;  
 RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,  
 RA Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;  
 RT "Comparative genomics of *Salmonella enterica* serovar Typhi strains Ty2 and CT18.";  
 RRL J. Bacteriol. 185:2330-2337(2003).  
 CC |- SIMILARITY: TO THE N TERMINAL PART OF THE UVRC FAMILY.  
 CC  
 CC This SWISS-PROT entry is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - nucleotide excision repair. Incises the DNA farther away from the lesion than uvrc. Not able to incise the 5' site of a lesion. When a lesion remains because uvrc is not able to induce the 3' incision. The combined action of cho and uvrc broadens the substrate range of nucleotide excision repair (By similarity).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).  
 CC  
 CC This SWISS-PROT entry is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - nucleotide excision repair. Incises the DNA farther away from the lesion than uvrc. Not able to incise the 5' site of a lesion. When a lesion remains because uvrc is not able to induce the 3' incision. The combined action of cho and uvrc broadens the substrate range of nucleotide excision repair (By similarity).  
 CC  
 CC |- SIMILARITY: TO THE N TERMINAL PART OF THE UVRC FAMILY.  
 CC  
 CC This SWISS-PROT entry is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).  
 CC  
 DR EMBL: AL672771; CAD02045; ALT\_INIT.  
 DR IntPro: AE016838; AA068844; ALT\_INIT.  
 DR InterPro: IPR003035; Uvrc\_N.  
 DR Pfam: PF01541; Exc\_endo\_N.  
 DR SMART: SM00465; GIY\_C\_1.  
 DR PROSITE: PS50164; UVRC\_1.  
 DR SOS\_RESPONSE\_HYDROLASE; Excision\_nuclease; DNA\_repair;  
 DR DNA\_recombination; DNA\_excision; Complete\_proteome.  
 DR SEQUENCE: 33332 MW; 4985505B856F7EE0 CR064;  
 SQ 293 AA;  
 DR Query Match Score 54; DB 1; Length 293;  
 Best Local Similarity 50.0%; Pred. No. 6;  
 Matches 6; Mismatches 5; Indels 0; Gaps 0;  
 DE Protein  
 DE Endonuclease cho (EC 3.1.25.-) (Endonuclease cho)  
 DE (Uvrc homolog  
 DE protein).  
 GN CHO OR 22771 OR ECS2447.  
 OC Escherichia coli O157:H7.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Escherichia.  
 OC NCBI\_TaxID=83334;  
 RN [1] SEQUENCE FROM N.A.  
 RP RC STRAIN=C0157:H7 / EDL933 / ATCC 700927;  
 RC MEDLINE=21074935; PubMed=11206551;  
 RX Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
 RA Posfai G., Hackert J., Klink S., Boutin A., Shao Y., Miller L.,  
 RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamitis K.,  
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
 OS *Salmonella typhimurium*, *Enterobacteriaceae*, *Gammaproteobacteria*, *Enterobacteriales*,  
 OS *Escherichia coli* O157:H7.



vitro, the incision activity of cho is uvrA and uvrB dependent. When a lesion remains because uvrC is not able to induce the 3' incision, cho incises the DNA. Then uvrC makes the 5' incision. The combined action of cho and uvrC broadens the substrate range of nucleotide excision repair.

-!- INDUCTION: Repressed by lexA.

-!- SIMILARITY: TO THE N-TERMINAL PART OF THE UVRC FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; AE00269; AAC74811.1; -.  
 DR PIR; B64933; E64933;  
 DR Ecogene; EG13933; cho.  
 DR InterPro; IPR000305; Uvrc\_N.  
 DR Pfam; PF01541; Exc1\_endo\_N\_1.  
 DR SMART; SM00465; GIYCY\_1.  
 DR PROSITE; PS50164; UVRC\_1.  
 KW SOS response; Hydrolase; DNA excision nuclease; DNA repair;  
 KW DNA recombination; DNA excision; Complete proteome.  
 SQ SEQUENCE 295 AA; 33739 MW; DCEFAADD855EPF327 CRC64;

Query Match Score 54; DB 1; Length 295;  
 Best Local Similarity 50.0%; Pred. No. 6;  
 Matches 11; Conservative 6; Mismatches 5; Indels 0; Gaps 0;  
 SEQUENCE 295 AA; 33739 MW; DCEFAADD855EPF327 CRC64;

Qy 10 LSAQLIQARLMKBEESPVVSWRL 31  
 Db 89 IGALLIBARLIKEQQPLFNKRL 110

RESULT 15  
 CHO\_SHIFL STANDARD PRT; 295 AA.  
 AC P59361;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DE Excinuclease cho (EC 3.1.25.-) (Endonuclease cho) (Uvrc homolog protein).  
 DE protein).  
 GN CHO OR SP1485 OR S1602.  
 OC Shigella flexneri.  
 OC Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Shigellales.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=301 / Serotype 2a;  
 RX MEDLINE=22272406; PubMed=12384590;  
 RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,  
 RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,  
 RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,  
 RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,  
 RA Yu J.;  
 RT "Genome sequence of *Shigella flexneri* 2a: insights into pathogenicity through comparison with genomes of *Escherichia coli* K12 and O157.";  
 RL Nucleic Acids Res. 30:4432-4441(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=2457T / ATCC 700930 / Serotype 2a;  
 RX MEDLINE=22580274; PubMed=12704152;  
 RA Fournier G., Goldberg M.B., Burland V., Venkatesan M.M., Deng W., Darlington A.,  
 RA Mau B., Payne N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,  
 RA Schwartz D.C., Blattner F.R.;  
 RT "Complete genome sequence and comparative genomics of *Shigella flexneri* serotype 2a strain 2457T.";  
 RT Infect. Immun. 71:2775-2786 (2003).

CC -!- FUNCTION: Incises the DNA at the 3' side of a lesion during nucleotide excision repair. Incises the DNA farther away from the lesion than uvrC. Not able to incise the 5' site of a lesion. When a lesion remains because uvrC is not able to induce the 3' incision, cho incises the DNA. Then uvrC makes the 5' incision. The combined action of cho and uvrC broadens the substrate range of nucleotide excision repair (BY similarity).

CC -!- SIMILARITY: TO THE N-TERMINAL PART OF THE UVRC FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC DR EMBL; AE015171; ANN307.1; ALT\_INIT.  
 DR EMBL; AE016283; AAP1670.1; -.  
 DR InterPro; IPR000305; Uvrc\_N.  
 DR PF01541; Exc1\_endo\_N\_1.  
 DR SMART; SM00465; GIYCY\_1.  
 DR PROSITE; PS50164; UVRC\_1.  
 DR SOS response; Hydrolase; Excision nuclease; DNA repair;  
 KW DNA recombination; DNA excision; Complete proteome.  
 SQ SEQUENCE 295 AA; 33705 MW; 875D7DF594BF472 CRC64;

Query Match Score 54; DB 1; Length 295;  
 Best Local Similarity 50.0%; Pred. No. 6;  
 Matches 11; Conservative 6; Mismatches 5; Indels 0; Gaps 0;  
 SEQUENCE 295 AA; 33705 MW; 875D7DF594BF472 CRC64;

Qy 10 LSAQLIQARLMKBEESPVVSWRL 31  
 Db 89 IGALLIBARLIKEQQPLFNKRL 110

Search completed: July 7, 2004, 17:08:12  
 Job time : 18 secs

This Page Blank (uspto)